DOCUMENT RESUME

ED 407 957 JC 970 284

AUTHOR Georgakakos, John H.
TITLE Using the California Chemistry Diagnostic Test and Other Student Background Factors To Predict Grades and Success in General Chemistry I.
INSTITUTION Riverside Community Coll., CA.
PUB DATE May 97
NOTE 110p.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Chemistry; Community Colleges; Diagnostic Tests; *Grades (Scholastic); Predictive Validity; *Predictor Variables; Regression (Statistics); Science Instruction; *Student Characteristics; Success; Two Year College Students; Two Year Colleges

ABSTRACT

A study was undertaken at California's Riverside Community College District to determine predictors of grades and success rates in a first-semester general chemistry course. Fall 1996 course outcomes were analyzed for two groups of students: 137 who had completed the course, successfully or not, and 178 who had dropped it. Data were also collected for the sample from results of the California Chemistry Diagnostic Test (CCDT) and a pretest survey, both administered on the second day of the semester, and District records of students' grades and characteristics. From these data, 19 independent variables were identified and a regression analysis was run to determine which factors were associated with student success. The analysis indicated that grades were predicted by a 4-variable model including students' cumulative college grade point average (GPA), CCDT score, "B" or better grades received in advanced science and mathematics courses at high school, and Asian ethnicity. Course success was found to be related to college GPA; CCDT score; receiving a "B" or better in preparatory chemistry; and "B" or better grades in high school algebra, geometry, and biology. Appendixes provide the pretest instrument; a list of independent variables; and discussions of data techniques related to missing values, frequencies, variable reduction and screening, correlations and basic statistics, analysis of models, and comparisons of general chemistry results. Contains 42 references and 16 tables and figures. (HAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 407 957

Using the California Chemistry Diagnostic Test and Other Student Background Factors to Predict Grades and Success in General Chemistry I

by

John H. Georgakakos, Ph.D.

Associate Professor of Chemistry

Riverside Community College

May, 1997

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. H. Georgakakos

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

JC 970 284

BEST COPY AVAILABLE

ACKNOWLEDGEMENTS

Many people have helped to make this research endeavor a success. I extend my appreciation to the Riverside Community College District Board of Trustees and President Dr. Salvatore Rotella for their commitment to the spirit academic inquiry, as evidenced by the many faculty members such as myself whose research efforts they have openly encouraged and supported. I am grateful to Tom Johnson—RCC's Vice President of Academic Affairs—for helping me integrate my research plans into a finished sabbatical leave request and to the members of the Professional Growth & Sabbatical Leave Committee for their support and approval of the request. To my colleagues in the Chemistry Department—Mike Amrich, Doug Bond, Wilhelmina Hathaway, Joan Pleasants, Diane Marsh, Ellen Kime-Hunt, Sariwan Tjandra, and J. P. Hopper—thank you for allowing me to come into your classes to collect the necessary survey/diagnostic test data. To Dianna Jones, User Support Coordinator from Computing Services, and Arlene Johnson, Instructional Department Specialist for the Sciences, thank you for providing the necessary District-wide background and grades data on participating students. I am especially indebted to Jeanie Nishime, Dean of Counseling & Student Support Services, for helping me obtain the needed number of California Chemistry Diagnostic Tests. I also wish to thank UCLA's Dr. Arlene Russell for her insights and perspectives on the CCDT. Last but not least comes a special tip of the hat to the Chemistry 1A and Chemistry 3 students who participated in this project. It is through their participation that the results of the present study might be applied toward helping their unknown colleagues yet to come.

ABSTRACT

In order to determine which variables would be significant predictors of grades and success (C grade or better) in a first-semester general chemistry course, data collected on a sample of Riverside Community College District students who attempted General Chemistry I in the fall of 1996 were subjected to stepwise multiple regression analyses. Respective sample N's for the "grades" and "success" regressions were 137 and 178. The latter sample was 54% Caucasian, 14% Hispanic, 5% Black, 19% Asian/Pacific Is., 52% male, and 75% under 26 years of age. Various data reduction techniques, including principal components analysis, led to 19 independent variables (possible predictors) which were submitted to regression: student scores on the California Chemistry Diagnostic Test (CCDT); age; 4 dichotomous ethnicity variables (ASIAN, BLACK, HISPANIC, WHITE); sum of college courses previously taken in: (1) calculus and physics, (2) biology, geometry/trigonometry, algebra, and physical science; expected course grade; high school grade point average (GPA); whether the student took the course at night; gender; total self-rated ability scores (quantitative & verbal); whether the student had achieved a B or better in: (1) Introductory College Chemistry (Chemistry 2A), (2) Fundamentals of College Chemistry (Chemistry 3); whether the student had a cumulative college GPA $\geq$ 3.0; and two high school science/mathematics background factors. Two 4-variable models resulted: Grades were predicted by having a cumulative college GPA $\geq$ 3.0 (standardized regression coefficient = .30); CCDT score (.21); a high school background factor (.22), which represented the total number of more-

advanced high school science and mathematics courses—chemistry, physics, calculus, trigonometry—for which the student had achieved a B or higher; and being Asian (-.20). The variables explained 23% of chemistry-grade variance. Logistic regression showed success was predicted by college GPA ≥ 3.0 (.66); CCDT score (.57); having achieved a B or better in Chemistry 3 (.45)(a preparatory course for general chemistry); and the second high school background factor (.36), which represented the sum score of B-or-better grades in Algebras 1 & 2, geometry, and biology. The logit model (based on a cut-score probability of success of .5) was able to classify 71% of the entire sample and 89% of the successful students correctly. A separate dichotomous linear regression yielded the same 4-variable model and classification table as the logistic regression model and explained virtually the same amount of error variance in the criterion (16%). Model reliabilities are supported by collinearity diagnostics, residual analyses, cross-validation studies, and the trended stepwise behaviors of Mallow's C_p and the adjusted R^2. For advisement and placement the logit equation is more to be recommended since its predictions were associated with little disproportionate impact. The CCDT mean for a sample of 57 Chemistry 3 students who were also administered the test at the beginning of the term was 14.8, significantly smaller than the 20.1 value observed for the General Chemistry I sample.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	i
ABSTRACT	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
INTRODUCTION	1
Background and General Problem Statement	1
Multiple Measures	2
RELATED LITERATURE	3
METHODOLOGY	9
Subjects	9
Preliminary Background: Site of the Study	9
Subjects and Samples	9
Instruments and Variables	11
Experimental Procedures and Data Collection	14
ANALYSIS OF THE DATA AND RESULTS	14
Introduction and Statistical Guidelines	14
Basic Data-handling Strategy	15
Assumptions of Linear Regression	15
Strategy for Selecting an Optimum Regression Equation	16
Model Reliability and Validity	17
Linear Regression Results for the Prediction of Grades	19
Reliability/Validity Concerns	23
Results for the Prediction of Success	25
Logistic Regression Equation and Variables	25
Comparison of Logistic Regression with Linear Regression	28
Classification Tables for Success	30
Comparison of CCDT Scores between the General Chemistry and Fundamentals of Chemistry Courses	33

DISCUSSION 36
Disproportionate Impact 36
Conclusions and Recommendations 37
APPENDICES
I. 1996 STUDENT INFORMATION FORM AI-1
II. VARIABLES AII-1
III. MISSING VALUES, FREQUENCIES, and OTHER DATA HOUSEKEEPING AIII-1
Majors AIII-5
IV. VARIABLE REDUCTION and SCREENING REGRESSORS AIV-1
Data Reduction Using Principal Component Analysis AIV-1
Self-rated Ability Scales AIV-2
Background in College Science and Mathematics AIV-6
High School Course Grades AIV-8
Elimination of Other Variables AIV-11
Final List of Regression Variables AIV-12
SPSS Regression Programs AIV-13
V. CORRELATIONS AND BASIC STATISTICS AV-1
VI. ADJUSTED R^2 AND MALLOW'S Cp AVI-1
VII. ANALYSIS OF MODELS: SPLIT SAMPLES AND CROSS-VALIDATION AVII-1
Introduction AVII-1
Split-sample Analyses on N = 137 Group: Chemistry Grades AVII-2
Split-sample Analyses on N = 178 Group: Success AVII-4
VIII. ANALYSIS OF MODELS: RESIDUALS AVIII-1
Background on Residuals AVIII-1
Testing Residuals for Outliers AVIII-2
Linear Regressions AVIII-3
Grades as Dependent Variable (N = 137) AVIII-4
Linear Regresion: SUCC as Dependent Variable (N = 178) AVIII-6
Logistic Regression Diagnostics (N = 178) AVIII-11

IX. SOME COMPARISONS OF GENERAL CHEMISTRY RESULTS IX-1
CCDT Score Means and General Chemistry Course Grades IX-1
10 Factor Comparison of Successful with Unsuccessful Ss IX-3
REFERENCES R-1

LIST OF TABLES

Table		Page
1	Stepwise Regression Results for Predicting General Chemistry Grades	21
2	Prediction of Success in General Chemistry (N = 178): Coefficients and Fractional Weightings for Logistic and Linear Regression	28
3	Classification of General Chemistry Students into Unsuccessful (0) or Successful (1) Categories of Course Completion from Logistic Regression with College GPA ≥ 3.0, CCDT Scores, Chemistry 3 Grade B or Better, and High School Background Factor 2 as Response Variables	31
4	Actual Chances and Integrated Odds of Being Successful in General Chemistry for P_S Values Equal to or Greater than That Shown	33
5	Correlations between CCDT Score and Course Grade by Gender and Ethnic Groups for Chemistry 1A and Chemistry 3	34
6	CCDT Means and Standard Deviations by Gender and Ethnic Groups for Chemistry 1A and Chemistry 3	35
7	Correlations of Predicted Outcomes (Grades and Success) with Demographic Variables Related to Age, Gender, and Ethnic Grouping	36
AII-1	Listing of Variables	AII-2
AIV-1	Correlations: Self-rated Abilities for N = 178 Sample (below diagonal) and N = 137 Sample (above diagonal)	AIV-3

LIST OF TABLES—continued

Table		Page
AV-1	Basic Statistics: Chemistry Grades, Success and 19 Potential Predictors	AV-1
AV-2	Correlations: Chemistry Grades, Success, and Independent Variables for N = 178 Sample (**below diagonal**) and N = 137 Sample (**above diagonal**)	AV-2
AVII-1	Comparison of Final Step Regression Constants and Standard Errors	AVII-4
AVIII-1	Basic Statistics for Residual Differences from Different Regression Models	AVIII-13
AIX-1	ANOVA Table for Chemistry 1A CCDT Scores with Chemistry Grades as Factor	AIX-1
AIX-2	Pairwise Contrasts of CCDT Score Means by General Chemistry Course Grade	AIX-2
AIX-3	Multiple Variable Contrasting: Unsuccessful with Successful Categories of General Chemistry using the Entire Sample for Reference	AIX-4

LIST OF FIGURES

Figure		Page
1	Observed Groups and Predicted Probabilities	31
AIII-1	Histogram of Academic Ability	AIII-2
AIII-2	Histogram of Student Scores on the California Chemistry Diagnostic Test	AIII-4
AIII-3	Histogram of Grades in General Chemistry	AIII-5
AIV-1	Principal Components for Self-rated Abilities	AIV-5
AIV-2	Principal Components for Self-rated Abilities Using 1991 College-wide Student Data	AIV-5
AIV-3	Principal Components for College Courses Taken	AIV-7
AIV-4	High School Course Background Factors 1 & 2	AIV-10
AVIII-1	Histogram of Standardized Residuals	AVIII-4
AVIII-2	Normal Plot of Regression Standardized Residuals	AVIII-5
AVIII-3	Z_i vs. SYNGPAB for Chemistry-grades Sample	AVIII-7
AVIII-4	Z_i vs. CCDT Scores for Chemistry-grades Sample	AVIII-7
AVIII-5	Partial Regression Plot: Partial Residuals of CHMGRD vs. Partial Residuals of SYNGPAB	AVIII-8
AVIII-6	Distribution of Standardized Residuals	AVIII-9
AVIII-7	Histogram of Integerized Residuals	AVIII-10
AVIII-8	Residual Differences for Logistic Regression	AVIII-12

INTRODUCTION

Background and General Problem Statement

Riverside Community College District (RCCD) students who wish to enroll in General Chemistry I (Chemistry 1A) but who have not yet taken one of the course's approved prerequisites—e.g., high school chemistry, introductory college chemistry (Chemistry 2A), or a fundamentals of college chemistry course (Chemistry 3)—may fill out an appeals petition and thereby challenge the prerequisite on the basis of having a knowledge or ability to succeed in the course despite not meeting the prerequisite. It is the Chemistry Department Chair who will "review the petition and provide written notification to the student within five (5) working days" (RCCD, 1994). The Chair's decision to approve or deny the petition is based, for the most part, on an interview with the student. Absent from this process, however, are objective assessment data that speak directly to the student's chances for passing the course if allowed to enroll in it. The present research was designed to establish and validate regression equations which could be used to predict grades and success rates in first semester general chemistry for these petitioning students. It is anticipated that the inclusion of objective forecasts of student outcomes in the deliberation process will allow for more refined and accurate decisions (greater chance for successful calls) while at the same time improve the precision (less variance) in the process itself, with the student being the ultimate beneficiary.

Multiple Measures

One instrument designed specifically to gauge a student's readiness in terms of knowledge, skills, and ability for being successful in a traditional general chemistry course is the recently developed California Chemistry Diagnostic Test—CCDT—(Russell, 1994). It has been shown to be a valid predictor of performance for students in all three of California's higher education systems—University (UC), State University and Colleges (CSU), and Community Colleges (CCC) (Russell, 1994; Karpp, 1995). Consequently, the CCDT was the assessment instrument of choice in this study. With student scores on the CCDT considered to be one viable predictor, others were sought—these based on certain student background characteristics and academic experiences finding support in the literature as to their ability to forecast achievement in science or chemistry (see next section).

Of course having more than one significant predictor has its merits, allowing the researcher to triangulate, as it were, on the outcomes, in this case grades and success. Statistically speaking, as more significant independent variables become added to a regression equation, the amount of uncertainty in describing the dependent variable tends to become less. In addition, Title V Matriculation Regulations (Board of Governors California Community Colleges, 1993) prohibits use of any single assessment instrument "by itself" for student placement, calling instead for multiple assessment measures which do not have high correlations with each other. Thus, it was hoped that those independent

variables which ultimately "made it" into final regression equations would not be highly intercorrelated. Although the present regressions for grades and success utilized 19 potential predictor variables, only 4 (CCDT scores and 3 others) were found to be significant predictors in each of the two final regression equations (see Results).

RELATED LITERATURE

A number of studies have examined the roll of various student background characteristics and their ability to influence student achievement in chemistry. A summary of the findings from the early 1950s to the mid 1980s may be found in Nordstrom (1989) along with a listing of 16 related references from that period. According to Nordstrom, the most commonly reported predictors of performance up to that time were: Scholastic Aptitude Test (SAT) scores, American College Testing Program (ACT) scores, grades earned in high school chemistry and mathematics, high school grade point average (GPA) or rank, and locally developed placement examinations. The overall picture today looks much as it did a decade or so ago, although as will be shown below, new perspectives from the literature have come along and added some luster.

Boujaoude and Giuliano (1991) addressed the question: To what extent does prior knowledge, reasoning ability, attitude, and gender affect achievement in chemistry (as measured by a course final exam)? Subjects (Ss) were 114 females and 85 males enrolled in a second

semester non-science freshman chemistry course in a New York State private university. The results support a conclusion that greater chemistry achievement is predicted for students who have more subject prior knowledge, are at higher logical reasoning levels, and have an attitude toward the subject that involves intrinsic motivation to study and a desire to grasp the deeper meaning of the material such as how ideas and topics relate to one another (as opposed to a reproducing orientation). No sex differences were found. These findings dovetail nicely with those of Chandran, Treagust, and Tobin (1987) who found student's formal reasoning ability and prior knowledge in chemistry to be significant predictors of achievement of grade 11 high school students.

Georgakakos (1995) used a multiple regression approach to uncover factors influencing the science reasoning (SR) ability of 190 community college Ss who were enrolled in at least one freshman or sophomore level college chemistry course and who were pre- and posttested at the beginning and end of the semester with the ACT College Assessment of Academic Proficiency (CAAP) SR Test. After taking SR pretest scores into account, significant positive direct effects were observed for reading and mathematics ability (measured by the ACT ASSET Reading and Numerical Skills Tests), number of science units taken during the test semester, and for semester GPA in chemistry (the Pearson product-moment correlation coefficient [r] between the SR posttest and semester GPA in chemistry was .34, two-tailed $p < .01$). Although the dichotomous ethnicity variable WHITE also exhibited a positive direct effect and had a moderately high correlation of .45 with

the posttest, no ethnic group or gender had a clear advantage in chemistry GPA. However, significantly higher chemistry grades were achieved by those students who rated themselves highly at the beginning of the term in academic or science reasoning ability ($r = .21$).

Spencer (1996) has recently examined the relationship between student scores on the mathematical SAT test (SAT-M) and grades in both 1st and 2nd semester general chemistry at Oberlin College for N = 1,160 and 849 Ss who had taken these courses between academic 1987 and 1995. A clear linear relationship between chemistry grade and SAT-M scores was demonstrated for each course. Focusing on the results for the first semester course only (since these more clearly relate to the present study), one finds a SAT-M score of 500 predicts a "C," 600 a "C+," and 700 a "B." The authors report no significant gender difference in grades. Also evident is that chemistry and biochemistry majors earned significantly higher than average grades ("B+" vs. "B-"). (*According to Pascarella and Terenzini (1991), student major has its strongest impact in the content areas consistent with the major.*) Grades varied across the ethnicity groupings, and when converted to a numerical 4.0 scale, read approximately as follows for 1st semester general chemistry: Blacks 2.0, Latinos 2.4, Asians 2.5, Whites 2.8.

In an effort to examine the psychological factors that may undergird consistent gender and ethnic group differences found in science achievement, Garcia, Yu, and Coppola (1993) performed a hierarchical regression analysis where first semester organic chemistry grades of 557 college students was the output variable. Adding

demographic variables in the first stage revealed negative standard regression coefficients (beta, β) of -.12 and -.13 for being female and for being African-American or Hispanic; however, these variables accounted for only 3% of the variance in grades and had their coefficients reduced to insignificance during the second stage of regression where high school GPA (β = .18), verbal SAT scores (β = .22), and scores on a chemistry placement test (β = .19) were taken into account. Additional grade variance (from 21 to 34%) was explained by including motivational factors—such as intrinsic goal orientation (β = .15) and test anxiety (β = -.17)—at stage 3 and learning strategies—rehearsal (β = -.09) and effort management (β = .19) at the fourth and final stage. The authors conclude academic preparedness, mastery orientation, and high levels of effort are associated with higher grades, while poorer performance is related to high test anxiety and use of rote rehearsal strategies. They suggest academic preparedness, motivation, and the use of learning strategies are important mediators in the achievement of woman and minorities in science. These declarations may be ushered next to the results of Johnson and Butts (1983) which underscore the proposition that the greater the amount of "engaged time"[1] spent by a student in trying to master the material presented in a science course the greater his or her level of achievement in the course will be. Neither, are they incompatible with the Razali and Yager (1994) survey data which show

---

[1] Note: Pace (1982) uses the term quality of student effort. Dickinson and O'Connell (1990) talk of quality of study. In Astin (1987) terms engaged time would be considered a measure of student involvement.

that college chemistry professors believe it is more important for incoming college students to possess certain abilities such as demonstrated study habits, careful note-taking ability, and ability to do algebra and higher mathematics than to possess specific knowledge of chemistry such as facts, definitions, and relationships.

This section closes by returning to the study by Nordstrom (1989) in which a discriminant analysis was performed on a sample of 980 engineering students who had taken a 5-credit, 1-semester accelerated general chemistry course between fall, 1980 and spring, 1989; the author sought to identify factors which could be used to predict successful ("A," "B," or "C") versus unsuccessful ("D," "F," or "W") performance. As an exploratory study, some 30 student academic background and demographic variables were investigated. Factors identified as the best predictors of performance in general chemistry—in order of standardized relative weightings—were the mathematics score on the student's college entrance examination (.48), high school GPA (.29), course grade in high school chemistry (.25), high school mathematics GPA (.24), and course grade in high school English (.20). The model correctly predicted the discriminant group for almost 3 out of 4 students (73.7%).

One intention of this section has been to underscore factors paramount as predictors of student performance in chemistry. Prior knowledge/academic experience is represented by grades and standardized test scores related to courses previously taken in chemistry, mathematics, or English. Certain personal traits and abilities

are captured by student self-ratings, college major, units taken, engaged time, study habits, motivation, and reasoning ability (and also, to some degree, by prior grades and standardized test scores[2]). Demographic variables were represented by ethnicity and gender. In short, the results of this section and those reported earlier for the CCDT serve to anchor the selection of independent variables employed and assayed in the present study.

---

[2] The following passage from Pascarella and Terenzini (1991) is noteworthy:

> A student's grades are probably the single most revealing indicator of his or her successful adjustment to the intellectual demands of a particular college's course of study. Without satisfactory grades, a student will not graduate from college, nor will he or she be admitted to graduate or professional school. Although heavily influenced by academic ability and intelligence, grades in college are not merely a function of those factors. Even with academic ability or intelligence taken into account, grades at the individual level are influenced by such factors as personal motivation, organization, study habits, and quality of effort. Thus, . . . grades tend to reflect not only requisite intellectual skills but also desirable work habits and attitudes (p. 388).

METHODOLOGY

Subjects

Preliminary Background: Site of the Study

The site of the study was the Riverside Community College District, a large comprehensive, multi-campus district of approximately 21,000 students. District students (98% of whom reside within the district's Southern California boundaries) may take classes at any of the three campuses—Riverside, Moreno Valley, Norco—where lower division, transfer-oriented curricula in the liberal arts and sciences along with a wide range of certified occupational programs and courses in continuing and developmental education are offered (RCCD, 1997). Briefly, profiling the District's students: 64% are continuing or returning students, 26% first-time students, about half attend classes during the day only, 54% are age 25 or younger, 60% are female, 52% Caucasian, 24% Hispanic, 12% African American, 9% Asian/Pacific Is. . Fall 1996 enrollment (not head count) was 22,865, with 63.9% at Riverside, 17.5% at Moreno Valley, and 18.6% at Norco, continuing a six-year trend that has seen a greater percentage of the total enrollment shift away from Riverside to the two newer campuses (RCCD Office of Planning and Development, 1996).

Subjects and Samples

Subjects for the study were RCCD students who attempted General Chemistry I (Chemistry 1A) or Fundamentals of Chemistry (Chemistry 3) during the fall semester of 1996 and who, during a single class session at the beginning of the first week of the semester,

completed a short student questionnaire (1996 Student Information Form) and then took the California Chemistry Diagnostic Test. Chemistry 1A Ss were divided into two overlapping but distinguishable sub-samples for the regression studies: The first group (N = 137) included Ss whose end-of-semester Chemistry 1A grade was "A," "B," "C," "D," or "F." The second, (N = 178) included the entire Chemistry 1A sample, that is, all members of the first group plus those who had dropped ("W") the course during the semester (no student received an incomplete [I] or a no credit [NC]). The Chemistry 3 sample (N = 57) was not used for regression purposes but rather as a comparison group for the general chemistry students with respect to CCDT scores.

A profile of the Chemistry 1A sample (N = 178) reveals 75% were ≤ 25 years of age, 48% were female, 54% were Caucasian, 14% Hispanic, 5% African American, 19% Asian/Pacific Is. . When these numbers are compared with the district-wide percentages, it is strikingly evident that the general chemistry sample is younger, has a larger percentage of Asians and a lower percentage of African Americans, Hispanics, and females. This is more typical than unusual. For example, an earlier RCCD study (Georgakakos, 1995) used a Fall 1991 Chemistry 1A sample of 91 students, where 78% were ≤ 25 years of age, 40% were female, 68% were Caucasian, 13% Hispanic, 6% African American, and 19% Asian/Pacific Is. . Although it becomes somewhat speculative to generalize between these two chemistry samples, it is fairly evident that their percentages are more nearly aligned with each other than with those for the overall student population. This claim is

even more robust when coupled with the fact that since 1991 the percentage of Caucasian students attending RCCD has decreased from 60% to 52%, while the Hispanic student population has increased from 18% to 24% (RCCD Office of Planning and Development, 1996). Nevertheless, without support from, say, a multivariate analysis of variance (MANOVA), such a comparison between samples across the barriers of time and space is less-than-complete.

Instruments and Variables

Variables in the study come from the following sources: 1) the California Chemistry Diagnostic Test (Form 1993), 2) a pretest survey instrument, 1996 Student Information Form , 3) RCCD computerized records.

The CCDT (available from the American Chemical Society [ACS] Division of Chemical Education) is a 45-minute, 44-item multiple choice examination designed to assess students in the following eight topic areas considered by the test developers as normally taught in a first-term general chemistry course: (1) compounds and elements, (2) states of matter, (3) reactions of matter, (4) structure of matter, (5) periodic properties, (6) solutions, (7) qualitative concepts in thermodynamics and kinetics, and (8) laboratory skills. The Kuder-Richardson Formula 20 (KR-20) reliability reported for community college examinees is .86 (Russell, 1994). Predictive validities for the CCDT vis-à-vis general chemistry performance are contained in Kaarp (1995) and in Russell (1994) and will be elaborated upon in a later section.

The 1996 Student Information Form (see Appendix I) was patterned after those used by Astin in his national surveys (Astin, Green, Korn, Schalit, & Berz, 1988). Students used a five-point Likert scale to rate themselves on 9 ability items: academic ability, general knowledge, problem solving ability, science reasoning ability, reading ability, writing ability, lecture notetaking ability, intellectual self-confidence, and social self-confidence. They also provided information related **to high school courses taken and grade received** in biology, chemistry, physics, physical science, Algebra I, Algebra II, geometry, trigonometry, and calculus; **to college courses and grade received in**: Introductory Chemistry (Chemistry 2A and Chemistry 2B) as well as Chemistry 3; **and to number of college courses completed in**: biological science, physics, physical science/geology, algebra, geometry and/or trigonometry, and calculus. They also provided estimates for: high school GPA, number of hours/week expected to be spent studying for the course, and expected course grade.

Test-retest statistics and reliabilities for items identical to or comparable with those of the present survey are available in detail elsewhere (Astin, Christian, & Henson, 1975; Boruch & Creager, 1972; Pace, 1982; Georgakakos, 1995). In these writings strong arguments are presented for the utility and reliability of most of the single-item questions typically encountered in surveys of college students. It is not the intent here to reconstruct the various arguments or positions of these researchers, but rather to present the general findings. Accordingly, items related to declaration of age, sex, education level, citizenship,

ethnicity (particularly the major groupings), and average high school grades tend to be most reliable, with r values in the .9-.95 range. Items asking for students' previous course experience, previous achievements or activities engaged-in, degree aspirations, goals, and college major show r values more between .8 and .9, provided redeclaration is done fairly soon, say, within a few weeks. Student reports about their own academic abilities and improvement (relative to their peers) also appear to be quite stable and credible. Less reliable, with an r of about .65, are items related to how students plan to spend their study hours.

District records, the remaining source for experimental data, included the student's: Chemistry 1A (or Chemistry 3) grade, ethnicity, gender, age, college major, educational goal, fall units enrolled in, job work hours per week, total hours attempted, total hours completed, and total hours passed as a District student.

The reliability of computer records of the kind discussed here must be considered close to unity in the sense that once a student has completed a course and received a grade, it is an unalterable fact, not subject to errors of faulty recall, deliberate distortion, item ambiguity, and the like which are indigenous to survey data. Granted some errors in record keeping can occur, but these tend to reduce reliabilities by only a few one-hundredths of a point (Boruch & Creager, 1972).

A listing and description of all the variables subjected to scrutiny of one kind or another during the course of this study is given in Appendix II.

Experimental Procedures and Data Collection

The student survey/diagnostic-testing phase of the experiment was completed by the close of the second day of the semester and included all Chemistry 1A and Chemistry 3 sections. Ss were instructed that their participation in the project would not affect their course grade but was crucial for helping the staff better serve the academic needs of future RCCD students. After being allowed 10 minutes to complete the 1996 Student Information Form, Ss were administered the 45-minute CCDT in accordance with the testing procedures specified by the ACS. For the remainder of the semester business in Chemistry 1A and Chemistry 3 proceeded more or less as usual. After the semester ended, District records containing student grades and other background factors mentioned earlier were combined with the survey/CCDT results forming a final data base which would be amenable to the statistical analyses described below. To help insure faculty and student anonymity, special numerical codes instead of names were used. The experimental period of the study is the entire fall semester of 1996.

ANALYSIS OF THE DATA AND RESULTS

Introduction and Statistical Guidelines

The pages which immediately follow are devoted to a layout of ground rules and general procedures fundamental for establishing the regression models.

Basic Data-handling Strategy:

1. Examine frequencies and normalities of all variables (see Appendix III).

2. Collect appropriate descriptive statistics on all variables.

3. Determine the correlation matrix for all candidate independent variables; include dependent variables as well.

4. Select a reasonable set of independent candidates to be submitted for regression analysis (Appendix IV).

5. Run the regression analysis and any other supplemental analyses appropriate to the group under discussion.

6. Test the reliability and validity of the obtained model.

7. Summarize and discuss the model and other related findings.

Assumptions of Linear Regression

The assumptions of linear regression are explicitly stated in Glass and Hopkins (1984, p. 141):

1. The Y scores are normally distributed at all points along the regression line; that is, the residuals are normally distributed. (There is no assumption that the independent variables are normally distributed.)

2. There is a linear relationship between the Y's and $\hat{Y}$'s (predicted Y's)—at all points along the straight regression line, the residuals have a mean of zero.

3. The variance of the residuals is homogeneous at all points along the regression line. This characteristic is known as homoscedasticity.

For statistically based conclusions to be valid, the statistical models involved must be reliable (since reliability is a necessary though not sufficient condition for validity), which in the case of regression means, among other things, being able to conform to the specifications above.

Strategy for Selecting an Optimum Regression Equation

Kleinbaum, Kupper, and Muller (1988, p.315) give five steps for selecting the "best" regression equation:

1. Specify the maximum model to be considered.
2. Specify a criterion for selecting a model.
3. Specify a strategy for applying the criterion.
4. Conduct the analysis.
5. Evaluate the reliability of the model chosen.

In terms of what constitutes the appropriate size for a model, arguments can be marshalled on behalf of large or small maximum models. Arguments for a large number of predictors remind researchers about the importance of including all possible predictors; of including interactions and higher-order power terms; about the importance of not omitting a predictor with a true non-zero regression coefficient, that is, of not committing a Type II (false negative) error; and about the advantage of incorporating as many statistical controls as possible. Arguments for small maximum models are based on parsimony; a predilection to focus on major predictors only; a desire to avoid Type I errors (false positives); a concern that possible collinearity problems will arise as more variables are considered; and a very practical concern for keeping the number of variables considered (k) appreciably smaller than the sample size (N).

Kleinbaum, Kupper, and Muller (1988) suggest the ratio of N to k should be ≥ 5.

The present study borrowed from both philosophies above. While a large number of input variables were explored, no quadratic or higher power terms were used for any variable, neither were any interaction terms considered. Great care was also taken to maintain a respectable ratio of sample size to variables. In addition, the stepwise multiple regression approach taken allowed for a systematic imposition of order on the data, which, according to Kleinbaum, Kupper, and Muller (1988), can increase reliability as well as the chance of finding a scientifically plausible model.

Afifi and Clark (1984) offer a number of criteria for variable selection in regression: Among stopping-rule options discussed are R^2, the adjusted R^2 ($\overline{R}^2$), Mallow's C_p, and the partial *F* test. While the partial *F* test was the criterion of choice for the regressions here, other stopping rules were not simply discounted since they offered alternative perspectives and were of great use in helping to ascertain the reliability of the various models. More detail regarding the use of C_p and $\overline{R}^2$ and their involvement in these analyses is given in Appendix VI.

Model Reliability and Validity

All regression models were held to four reliability criteria: 1) collinearity diagnostics (minimum tolerance and Mallow's test), 2) residual analysis, 3) cross-validation studies, and 4) examination of C_p and $\overline{R}^2$ as variables were added to the equation. In-depth coverage

of these procedures may be found in the following references: (SPSS, 1988, 1992a; Afifi & Clark, 1984; Kleinbaum, Kupper, & Muller, 1988). This notwithstanding, selected details about the application of these criteria to the present data are furnished in Appendices VI through VIII.

Questions of model validity center on whether inferences drawn from the model are appropriate and accurate (Crocker & Algina, 1986; McMillan & Schumacher, 1984). In terms of content, one would expect a high degree of face validity...if, say, chemistry grades were predicted by CCDT scores as well as by previous science course experience, academic ability, math skills, and the like, since such a result would seem intuitive and not be lacking for support from the research literature. In addition, yet in a similar vein, model validity is related to the validity of the instruments and variables employed in its construction (discussed in the "Methodology"). In terms of criterion-related validity, the question is: How justified and well-established is the statistical relationship of the outcome with the input variables? Some issues to consider here are the sample, sample size, restricted range in sample selection, criterion contamination, confounding, and predictive capability as well as reliability of the model. To assert that the samples chosen are more or less representative of RCCD students who would typically be involved with the curriculum under investigation appears to be somewhat justified. Furthermore, the Chemistry 1A sample is of sufficient size and broad enough background representation that restricted range is not a problem. Still, caution is advised in the ascription of the findings to other groups of students! Criterion contamination, alias "teaching to the

test," is a non-issue, since all predictive factors were in place before the third day of an eighteen-week semester; consequently, regular classroom activities had not yet begun. Now, confounding of variables can never be totally ruled out; however, the very nature of being able to reduce, strategically and lawfully, an initial slate of many candidate predictor variables (Appendix II) to a final roster of 19 requires a relatively complete knowledge of how all the original variables interrelate (see especially Appendices IV and V). Furthermore, since comparable counterparts to the likely predictors can be pointed to in the established literature, it is well to conclude that confounding of variables should pose no threat to the validity of the regression results. Predictive validities and reliabilities of each model are established and discussed in sections that follow.

Linear Regression Results for the Prediction of Grades

Application of the variable reduction procedures described in Appendix IV produced a final list of 19 independent variables that were subjected to a final regression run. The variables were CCDT scores (DTEST); student's age (AGE); dichotomous ethnicity variables (ASIAN, BLACK, HISPANIC, WHITE); sum of college courses previously taken in calculus and physics (CALPHYS#); sum of college courses previously taken in biology, geometry/trigonometry, algebra and physical science (BALGTPS#); expected chemistry grade (EXPGRD); high school GPA (HSGPA); whether the student took Chemistry 1A at night (NIGHT); being female (SEXFEM); overall score for self-rated quantitative (QUANTAB) and verbal ability (VERBAB); whether the student had

achieved a B or better in introductory college chemistry (PC2AGDB), or a fundamentals of college chemistry course (PC3GDB), or had a cumulative college GPA of 3.0 or higher (SYNGPAB); and two high school background variables derived from factor analysis which measured the number of advanced as well as fundamental high school science and mathematics courses in which a student earned B or better (HSBFACT1 and HSBFACT2, respectively). Basic statistics as well as the overall correlation matrix for these variables, including the outcome variable (CHMGRD), are given in Appendix V.

Computerized statistical procedures for the entire project were guided by SPSS (1992a, 1992b, 1992c) software. Both the linear and logistic regression programs used are displayed in Appendix IV, page AIV-13. The stepwise linear regression procedure allowed for selection of variables whose partial *F* test yielded a p value equal to or less than the designated alpha (α = .05) level of significance—denoted as PIN in the program (i.e. the probability of *F*-to-enter). Once selected, a variable remained in the equation unless its POUT (probability of *F*-to-remove) exceeded the maximum designated value of .10. These same p-value limits were assigned in the logistic regression program; except, instead of using *F* tests for entry and removal, the logit procedure used the score statistic and the conditional likelihood ratio test, respectively. These statistics are based on the chi-square distribution (see SPSS, 1992b; Kleinbaum, Kupper, & Muller, 1988). For large samples (such as employed here), the logistic regression p values—which stem from the use of maximum-likelihood-informed chi-square statistics—are

comparable to their linear regression counterparts—which stem from the use of least-squares-based *F* statistics (Afifi & Clark, 1984; Glass & Hopkins, 1984; Kleinbaum, Kupper, & Muller, 1988).

A summary of the stepwise regression for prediction of grades is given in Table 1. The first column displays the regression step number; the second, the variable which was entered at that step. Column three indicates the multiple R at each step, and the adjacent column on the right gives the simple r between the criterion and each of the entering independent variables. The remaining columns contain the standardized regression coefficients for the variables at each step of the analysis. After step 1, only SYNGPAB has "made it" into the equation; the numbers shown for the other variables are "beta-in" values, that is, the betas these variables would have <u>if</u> entered on the next step. So, when CCDT scores enters on step 2, its β is .27 (the dotted line helps one to visualize this natural progression from one step to the next).

Table 1

Stepwise Regression Results for Predicting General Chemistry Grades

Step#	Variable Entering	R	r	Beta[a] after Step 1	2	3	4
1	SYNGPAB	.296	.296	.30	.30	.31	.30
2	DTEST	.402	.273	.27	.27	.24	.21
3	HSBFACT1	.443	.212	.23	.19	.19	.22
4	ASIAN	.483	-.206	-.18	-.16	-.20	-.20

<u>Note</u>. N = 137.

[a] The coefficient for any variable not yet in the equation shows the beta that variable would receive if it were entered into the equation on the next step.

In general absolute values for β should decrease going down a column as less potent predictors are taken into account. They should also decrease in going from left to right across the table as more variables become added to the equation and share explained variance with those already aboard. The final multiple R of .483 indicates that 23.3% of the variance in chemistry grades is accounted for by the four predictors. Examination of the final coefficients from the step 4 reveals that having a cumulative college GPA of 3.0 or higher has about 1.5 times the predictive power as any of the other variables. Also evident for this sample is the fact that being Asian generally forecasts lower general chemistry grades. Also note when the ethnicity variable is entered with a beta of -.20 at step 4, the beta for high school background increases from .19 to .22, indication of a suppressor effect, that is, because Asians achieved at higher levels in high school chemistry, calculus, physics, and trigonometry ($r_{ASIAN,HSBACT1}$ = .16) than did their peers—most of whom were white ($r_{WHITE,HSBACT1}$ = -.15)—then controlling for being Asian at step 4 serves to augment the influence of having higher scores on the high school science/math background variable. To further highlight this flip-flopping of the ASIAN and WHITE variables ... if ASIAN is removed from the original roster of 19 variables, then WHITE enters the equation at step 4 with a positive beta of .19; this occurs basically because: of the 110 non-Asian Ss, 75 were Caucasian. If now, WHITE is also removed from the original roster, then no other variable enters, in particular, no other ethnic variable—BLACK or HISPANIC—this primarily because the latter two ethnic groups represented only small fractions of

the sample—3% and 12%, respectively. Of the four ethnicity-group variables, only WHITE had a positive correlation with chemistry grades (r = .21) compared with ASIAN (-.21), BLACK (-.20), and HISPANIC (-.07).

In terms of unstandardized regression coefficients (b), the final equation is:

$$\text{CHMGRD} = 1.1792 + .7096 \times \text{SYNGPAB} + .0379 \times \text{DTEST} + .2146 \times \text{HSBFACT1} - .5759 \times \text{ASIAN}$$

Equ(1).

Reliability/Validity Concerns

Recall the four model reliability criteria: 1) collinearity diagnostics, 2) residual analysis, 3) cross-validation, and 4) examination of C_p and $\overline{R}^2$ as variables are added to the equation. The last criterion, discussed more fully in Appendix VI, was used to confirm that variable selection in producing Equation (1) was indeed parsimonious; Mallow's C_p always decreased while R^2 and $\overline{R}^2$ each increased as variables were added.

Several tests for multicollinearity were used: First, the correlation matrix of the 19 submitted variables showed only moderate to low correlation coefficients. Second, the collinearity diagnostics in the SPSS Mallow's C_p procedure flagged no problems. Third, during regression analysis, SPSS calculates, at each step, for each potential predictor, a diagnostic known as the minimum tolerance. A minimum tolerance of 0

may be taken to mean a variable's effect is completely explicable in terms of other variables already in the equation; a value of 1 means that if the given variable is added to the equation, the effect carried by each predictor is unique, unrelated to the effects carried by other predictors. The alert for collinearity, then, is small tolerance values ... ≤ .1. None approached this threshold; the lowest minimum tolerance was .65 for WHITE when ASIAN entered at step 4.

Even though the split-sample-analysis results in Appendix VII only partially affirm the full model's reliability, the residual analysis data of Appendix VIII line up on the side of C_p, $\overline{R}^2$, and tolerance data in lending support. In sum, the model passes the tests of reliability. Other fundamentals pertinent to the validity of the model, such as criterion contamination and various sampling parameters, have also been addressed.

In terms of predictive validity: After rounding predicted values of Equation (1) to nearest whole numbers, each subject was assigned a letter grade. Comparison of these grades with actual grades earned yielded correct predictions for 37% of the cases, underpredictions for 34%, and overpredictions for 29%. At first glance these findings may not appear very promising, but a second look leaves more room for optimism. For example, of the 47 grades underpredicted, 32 were for students who had earned A—75% of these Ss received a predicted grade of B, and the rest C. Of the 39 Ss whose grade was overpredicted, 18 were C's who were assigned B's. Thus, while the equation may mispredict grades, it tends not to mispredict by more than

a single grade-point level and also seems able to predict whether a student will pass or not, even when the predicted and achieved grades are not matched exactly. Perhaps more feasible, tractable, and pragmatic is not to be overly concerned about forecasting exact grades but rather to focus on establishing predictable levels for passing, an objective more in keeping with results offered by a discriminant or two-group regression analysis. Recall Nordstrom's (1989) discriminant analysis where approximately 75% of the students were correctly classified. Prior to this result, the author reported a multiple R of .428 for a regression equation where college chemistry grades were predicted from high school GPA, SAT/ACT mathematics score, and high school chemistry grade. In contrast to the discriminant function, the regression equation yielded correct prediction for only 32.5% of the sample.

Results for the Prediction of Success

With success in General Chemistry I as the dependent variable, the 19 independent variables used earlier were subjected to both a logistic regression and a dichotomous linear regression analysis. Before any substantive comparisons between the equations developed using the two regression methods are attempted, a certain coverage must be given to some of the logistic regression results. The section can then close with a presentation of the logit-derived classification tables.

Logistic Regression Equation and Variables

The stepwise logistic regression algorithm led to a four-variable model. The logit equation is:

$$Z = .0862 \times DTEST + .2929 \times HSBFACT2 + 1.4857 \times PC3GDB + 1.3764 \times SYNGPAB - .3.7403$$ Equ(2),

where Z equals $\ln\left(\frac{P_S}{1-P_S}\right)$, the natural logarithm of the "odds" of being successful in General Chemistry I; P_S and $1\text{-}P_S$ are the probabilities of successful and unsuccessful course completion. From Equation (2) it is clear that higher CCDT scores and having a cumulative college GPA $\geq$ 3.0 are again significant positive factors. That HSBFACT2 and not HSBFACT1 appears in the equation could be a sign that <u>success</u> in general chemistry (as opposed to the attainment A's, B's and C's per se) is strategically more dependent upon achievement levels in fundamental high school science and mathematics courses, such as biology, algebra, and geometry, than upon achievement levels in more advanced high school science and mathematics courses. The statement should be regarded cautiously, however, since a slightly different sample was involved in arriving at the present regression result. Along this same line of reasoning the following question presented itself: Can one view the presence of the remaining positive predictor in Equation (2), PC3GDB (i.e., earned B or higher in Chemistry 3) as a partial supplantation of high school background factor 1? It seemed as though support for a "yes" answer to the question could be wrested from a few <u>positive</u> correlations: One, the relationship between Chemistry 3 grades

and high school chemistry grades, since scores on the displaced variable are by definition contingent upon achievement levels in high school chemistry. This correlation was indeed positive (.28) but not significant because only 19 students took both courses. Another monkey wrench here was the zero correlation between B-or-better variables HSCHMGDB and PC3GDB (N = 178).

A second correlation to examine was that between having taken chemistry in high school and having taken Chemistry 3: A common belief is the fundamentals college chemistry course is taken by students who lack high school chemistry experience—which would manifest as a negative correlation between the two variables. For the present sample, the "counter-intuitive" situation prevailed; r between the two course variables checked in at .1 but not significant. Still, 86% of the Ss who had taken Chemistry 3, had also taken chemistry in high school. The upshot of the off-road inquiry begun on the previous page is this: While only a feeble "yes" wrapped in uncertainty may be voiced to the question, it is rigidly clear that future research needs to address more comprehensively the relationship between students' involvement/grades in high school chemistry and their involvement/grades in a fundamentals college chemistry course.

Finally, although no ethnicity or other variables were significant, the overall correlation matrix involving all potential predictors as well as the criterion is arrayed in Appendix V for completeness.

Comparison of Logistic Regression with Linear Regression

The loglinear and linear regression computer outputs featured identical patterns of model development: same variables added at each step, none eliminated, regression halted after step 4. Standardized and unstandardized regression coefficients for the resultant equations are exhibited in Table 2. The loglinear β's were obtained by multiplying the model b's by the standard deviations of the associated variables. Fractional weightings ($\beta/\Sigma\beta$) were obtained as noted, by dividing the beta for the variable by the sum of the betas over all the predictors. As can be seen, each model explains virtually the same amount of variance in success (~16%).

Table 2

Prediction of Success in General Chemistry (N = 178): Coefficients and Fractional Weightings for Logistic and Linear Regression

		Loglinear Results			Linear Results		
Step#	Variable	b	β	$\beta/\Sigma\beta$	b	β	$\beta/\Sigma\beta$
1	SYNGPAB	1.376	.655	.32	.246	.246	.31
2	DTEST	.086	.573	.28	.016	.221	.28
3	PC3GDB	1.486	.449	.22	.262	.166	.21
4	HSBFACT2	.293	.357	.18	.060	.154	.20
	Constant	-3.740			-.185		
R^2			.15[a]			.16	

[a] The logistic regression procedure does not compute a multiple R. Aldrich and Nelson (1984) suggest an approximation to an R^2 by dividing the model chi-square by the sum of the model chi-square + N, which was what was done here.

Fractional weightings for each variable vary slightly, if at all, from one methodology to the next. Having a college GPA of 3.0 or higher shoulders 32% of the burden of prediction in the logit equation and 31% in the linear model. CCDT scores carries 28% regardless of method. Astin (1991, p.283) notes that provided the categorical split in the dependent variable is not too lopsided (e.g., ratios of 9 to 1 or greater), then the two regression procedures should lead to relatively similar findings. Here the unsuccessful/successful split is 34.3% to 65.7%.

In terms of reliability/validity concerns: The legitimacy of the four-variable model was corroborated by the step-over-step decrease in C_p with the accompanying step-over-step increase in $\overline{R}^2$ (Appendix VI). Collinearity tests also were all negative. (The smallest minimum tolerance was .71 for HSBFACT1 when HSBFACT2 entered at step 4; r between these two variables is .50.) Although logistic regression does not compute C_p, $\overline{R}^2$, and minimum tolerance values, basic conclusions reached from the use of these diagnostics regarding a particular categorical linear regression model also should apply to a companion logistic regression model ([i.e., one informed by the same correlation matrix] and [assuming pertinent variable-selection factors—such as, PIN, POUT, forward versus backward elimination, etc.—are the same]). Regarding the ability of the models to predict well for different samples, some confidence may be drawn from the average "shrinkage on cross-validation" value of .1 derived from ten different half-sample equations (Appendix VII). Lastly, residual analyses for both models pointed to no outliers (Appendix VIII). All told, the models for predicting success in

general chemistry are sustained by the reliability-assessment checklist. In fact, and in large part due to a larger initial sample size, the present categorical models appear to be even more robust than the continuous model used to predict chemistry grades.

Classification Tables for Success

Converting the logit equation to exponential form allows one to calculate P_S values for each student (P_{S_i} [also called P_S]) based on the values of his or her response variables; thus, $P_S = \left(\frac{1}{1+e^{-Z}}\right)$. Prior probabilities for successful and unsuccessful completion of Chemistry 1A are .66 and .34, respectively. In other words without use of any input variables, the odds of correctly classifying a student as successful are approximately 2 to 1. If on the other hand one uses P_S values for individuals as calculated from the logit equation, with the stipulation that those students whose P_S value is equal to or greater than .5 will be classified as prospectively successful and those whose P_S value is less than .5 will be classified as prospectively unsuccessful, one has the basis for the classification table presented on the next page. Notice that the prediction rate for the successful students is increased by 23 percentage points to 88.9%. However, the hit rate for classification of unsuccessful Ss is increased by only a few points to 36.1% (as attested to by the 39 Ss misclassified as "successful" in the upper right off-diagonal slot). Overall 70.8% of the Ss are correctly classified.

Table 3

Classification of General Chemistry Students into Unsuccessful (0) or Successful (1) Categories of Course Completion from Logistic Regression with College GPA ≥ 3.0, CCDT Scores, Chemistry 3 Grade B or Better, and High School Background Factor 2 as Response Variables

Predicted → Observed ↓	0	1	Percent Correct
0	22	39	36.07
1	13	104	88.89
		Overall	70.79

Note. This identical classification table was obtained, case for case, when the linear regression equation for success was used with the cut off for classifying Ss coincident with the logit cut level—that is, students whose predicted success score was equal to or greater than .5 were assigned to predicted group "1", while those whose predicted success was less than .5 were assigned to "0".

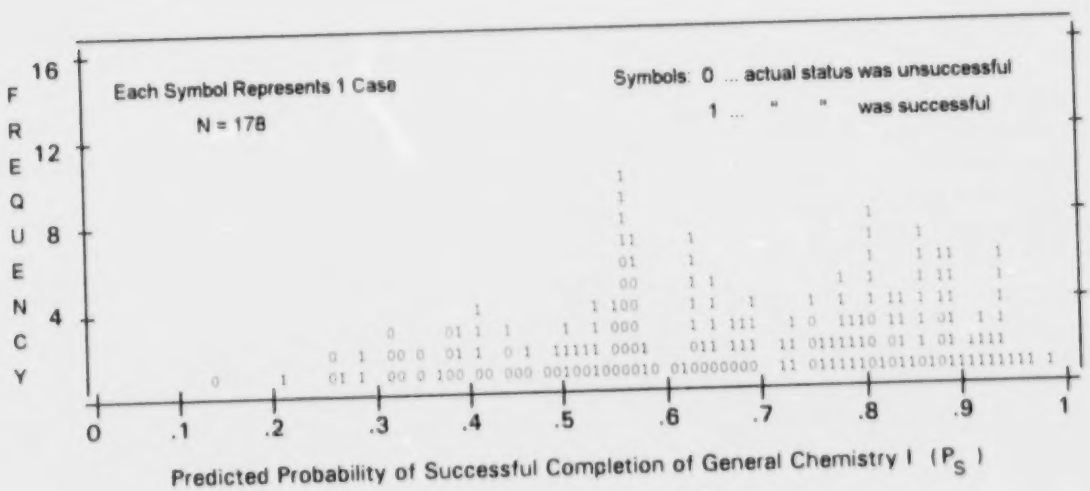

Figure 1 Observed Groups and Predicted Probabilities

More detail on these classifications is afforded by the histogram of estimated probabilities (Figure 1). As can be seen, almost half (18) of those misclassified as successful fall within the range of $.5 \leq P_S \leq .6$.

What is often done to stress the value of predictive equations such as Equation 2 is to compute P_S or odds for specified values of the input variables and compare these predicted outcomes to outcomes obtained using alternative input values: For example, consider two first-time college freshmen (no college GPA or Chemistry 3 experience) who wish to enter Chemistry 1A. Assume each student has a score of 4 on the high school background factor variable (e.g., B's in Algebras 1 & 2, geometry, and biology) but they differ in how well they performed on the diagnostic test. Let's say the first student attains the mean score of 20 (out of 44) on the test, while the second gets a 35. The calculated P_S and odds of passing for the first student are .63 and 1.70 to 1; for the second student these become .86 and 6.14 to 1. Notice that a 15 point increase in diagnostic-test performance has increased the odds of being successful by a factor of 3.6. Clearly, doing well on the CCDT greatly enhances the odds for success in general chemistry.

An alternate way of examining the results and which will also be useful in determining a student's prospects of success is to not just look at calculated probabilities but to in addition examine what these probabilities mean in terms of actual or "in-fact" success ratios for a particular group of students. For example, the histogram of estimated probabilities reveals that of the 80 Ss whose computed P_S value was equal to or greater than .7, 68 of these did successfully complete the

course while 12 did not; the in-fact chance of success for students in this cluster, then, is 68/80 or .85. Thus, while the computed odds of success at P_S = .70 are 2.3 to 1 the in-fact integrated odds of success for students whose P_S values equaled or exceeded .70 are 5.7 to 1. Some thinking along this line is summarized in Table 4.

Table 4

Actual Chances and Integrated Odds of Being Successful in General Chemistry for P_S Values Equal to or Greater than That Shown

	No P_S	$P_S \geq .5$	$P_S \geq .7$	$P_S \geq .85$	$P_S \geq .9$
Actual Chance of Success	.66	.73	.85	.89	1.00
Integrated Odds of Success	1.9	2.7	5.7	8.1	∞

Notice how dramatically the in-fact odds increase with increasing values of P_S. For any student in the sample whose calculated P_S reached .9, the chance of <u>not</u> completing the course successfully is zero%.

Comparison of CCDT Scores between the General Chemistry and Fundamentals of Chemistry Courses

Presented in Table 5 are correlations between CCDT score and course performance by gender and ethnic groups for the General Chemistry I and Fundamentals of Chemistry courses. Although course differences in the magnitudes of r for a given group may be observed,

none of these differences are statistically significant (as determined using the Fisher *Z*-transformation procedure described in Glass & Hopkins, 1984, p. 309).

Table 5

Correlations between CCDT Score and Course Grade by Gender and Ethnic Groups for Chemistry 1A and Chemistry 3

	Chemistry 1A		Chemistry 3	
Group	Correlation	N	Correlation	N
Males	.25*	73	.06	26
Females	.28*	64	.55*	23
Whites	.21	75	.34	30
Nonwhites	.27*	62	.20	19
Total	.27**	137	.32*	49

* signifies $p \leq .05$; ** signifies $p \leq .01$ (2-tailed).

The correlation coefficients confirm that the CCDT tends to predict chemistry grade for each group, particularly for Chemistry 1A. The total correlation of .27 for Chemistry 1A is a little smaller than the statewide value of .32 reported by Russell (1994) for a sample of 916 California Community College Ss.

That the test appears applicable to the prediction of performance in a fundamentals course, a purpose for which it was not specifically designed, is also of interest—although it should be noted that the correlation between SUCC and CCDT scores for the Chemistry 3 sample (N = 57) was only .01, compared with .25 for the Chemistry 1A (N = 178)

group. CCDT means and standard deviations for the N = 178 and N = 57 course samples are displayed in Table 6.

Table 6

CCDT Means and Standard Deviations by Gender and Ethnic Groups for Chemistry 1A and Chemistry 3

	Chemistry 1A			Chemistry 3		
Group	CCDT Mean	Std Dev	N	CCDT Mean	Std Dev	N
Males	20.90	6.40	92	15.63	4.36	27
Females	19.20	6.84	86	14.13	5.02	30
Asians	19.18	7.67	33	16.20	4.97	5
Blacks	17.63	6.74	8	12.00	5.35	4
Hispanics	18.71	6.34	24	11.78	3.67	9
Whites	21.39	6.33	96	15.66	4.78	35
Nonwhites	18.55	6.73	82	13.55	4.46	22
Total	20.08	6.65	178	14.84	4.74	57

As can be seen and as would be anticipated, Chemistry 1A CCDT means are greater than the Chemistry 3 means for all groups, and, of course, the differences are highly significant ($p < .001$) for the groups with the larger N's— males, females, whites, nonwhites, and total. For a given course the trend of means—whites higher than nonwhites, males higher than females—is similar to that reported by Kaarp (1995). For Kaarp these differences were both statistically significant, deriving from a total sample of 4,023 general chemistry participants; for the present general chemistry sample, statistical significance applies to the white-nonwhite but not to the male-female contrast. Russell (1994) reports a

CCDT mean of 22.37 for CCC students, a value which may be compared to 20.08 in the last row of the table.

DISCUSSION

Disproportionate Impact

A method of assessing the degree of disproportionate impact in the application of placement models is to correlate predicted scores with those variables which represent potentially affected groups. Table 7 presents such r values for variables associated with age, gender, and various ethnicities with predicted grade and predicted success. The large correlation of ASIAN with predicted grade shows a disproportionate impact, which is not surprising since the ASIAN variable is part of Equation 1. For predicted success, there is less room for concern since only a few percentage points of the variance in student success appears

Table 7

Correlations of Predicted Outcomes (Grades and Success) with Demographic Variables Related to Age, Gender, and Ethnic Grouping

Variable	Predicted Grade	Predicted Success
AGE	.139	.067
ASIAN	-.426**	-.046
BLACK	-.110	-.188*
HISPANIC	.074	.022
WHITE	.283**	.148*
SEXFEM	.011	.026

* signifies $p \leq .05$; ** signifies $p \leq .01$ (2-tailed).

attributable to an ethnicity variable (e.g., eta squared for WHITE is 2.2% and for BLACK 3.5%). No disproportionate impact due to age or gender is evidenced for either equation. Other contrasts and comparisons involving the final equation variables may be found in Appendix IX.

Conclusions and Recommendations

The statistical techniques and regression procedures discussed in the body and appendices of this paper have led to equations which may be employed to predict grades and chances of success in a first semester general chemistry course. The lack of collinear data, a normal distribution of residuals, an absence of outliers, the lawful selection of predictors, adequately parameterized models, a high degree of overlap between split-sample and full-sample equations, and reasonable average shrinkages on cross-validation all attest to the reliability of the models. Of the four variables found to predict grades, three ... having a cumulative college GPA $\geq$ 3.0, doing well on the CCDT, and having attained good grades in more advanced high school science/mathematics classes ... had positive betas. The fourth, being an ethnic minority, specifically Asian, was a negative predictor. Forecasting success also involved four predictors: CCDT score; college GPA $\geq$ 3.0; having earned B or better in high school algebra, geometry, biology; and having earned B or better in Chemistry 3—all with positive regression coefficients.

The establishment of cut scores and other methods for advising or placing students is a foremost dilemma that confronts many educators.

Judgment regarding whether students should be allowed to enroll in a particular course should be tempered by an examination of prediction data from more than one perspective. Equation 1 appears quite capable of reliable prediction of general chemistry grades, but the specter of disproportinate impact with whites advantaged over minorities, particularly over Asians, is evident. Still, its application should have value in terms of providing reasonable grade estimates for potential Chemistry 1A students. Equation 2 does not appear to be encumbered by the same ethnicity bias as does grade prediction; consequently, it is the equation of choice. Table IX-3 in Appendix IX contrasts means of pertinent demographic and academic achievement variables for the successful and unsuccessful chemistry Ss. The percentage of white Ss in the successful group averaged almost ten points higher than in the unsuccessful category, 57% to 48%; yet, the difference was not big enough to be statistically significant. What does account for the significant differences in student success are factors that involve readiness in chemistry and previous academic achievement. For example, 18% of the unsuccessful students reported having a college GPA of 3.0 or higher, compared to 43% for the successful group ($p = .001$). In addition, successful Ss had mean scores on high school background factors 1 & 2 that were 36% and 19% higher than their unsuccessful counterparts. They also significantly outperformed them on the CCDT (21.2 vs. 18.0, $p = .002$) and had a greater percentage of their membership with a B or better in a previously taken fundamentals college chemistry course (13% vs. 5%).

Overall, the classification accuracy of the loglinear model (with the P_S-cut-score level set equal to .5) was 71%. The model also classified 89% of the successful Ss correctly. While the predictive accuracy of the logit equation appears to be very good for P_S scores greater than .7 or less than .4, there still remains that in-between, borderline region where the likelihood for misclassifications is magnified. Consequently, great care must be taken in dealing with predicted scores which lie in this more nebulous, "gray" area.

A first recommendation of this paper, then, is that the present regression models be used to assist in the appropriate placement of students who desire to take general chemistry without first having taken certain established prerequisites.

A second recommendation, which serves as a corollary to the first, is to have the Chemistry Department Chair meet with the petitioning student to go over the prediction results. Every effort should be made to help the student understand how the results are being used to help direct him or her to the most appropriate chemistry course—i.e., the one most suited to the student's academic-knowledge level as well as academic need. The student should be apprised of exactly what a particular predicted score means in terms of probabilities and odds of passing the course. It is also envisioned that the student will be able to formulate an appreciation of the extent to which the calculated prediction and its various implications align with his or her own internal assessment of potential success in the course.

A third recommendation is that more comprehensive studies be undertaken which explore the relationship between students' recorded high school science/math grades and their involvement/performance in a fundamentals of college chemistry course ... also, the extent to which this relationship impacts various performance outcomes in general chemistry.

A fourth and final recommendation is that, in time, the present study should be updated using a new Chemistry 1A cohort: Future regression models may then be analyzed in the light of those developed here. For now, however, it suffices to hope that the added input furnished by CCDT scores, college GPA, and pertinent high school background variables will allow for more objective and accurate predictions about students' potential performance in general chemistry.

APPENDIX I

1996 STUDENT INFORMATION FORM

Dear Student:

The information in this form is being collected as part of a study of higher education at Riverside Community College. Your participation in this research is crucial in order that we may better serve the academic needs of our students. Your response will be held in the strictest professional confidence. You will ultimately be assigned a code number to insure your anonymity.

Sincerely,

Dr. John Georgakakos
Head Project Researcher

PLEASE ANSWER ALL QUESTIONS USING A #2 PENCIL

PLEASE PRINT (one letter or number per space)

NAME: _____ _____ _
LAST FIRST MI

SOCIAL SECURITY NUMBER: ___/__/___

1. Rate yourself on each of the following as compared with the average person your age. We want the most accurate estimate of how you see yourself.
(Mark one in each row)

	Highest 10%	Above Average	Average	Below Average	Lowest 10%
Academic ability	☐	☐	☐	☐	☐
General knowledge	☐	☐	☐	☐	☐
Problem solving ability	☐	☐	☐	☐	☐
Scientific reasoning ability	☐	☐	☐	☐	☐
Reading ability	☐	☐	☐	☐	☐
Writing ability	☐	☐	☐	☐	☐
Lecture notetaking ability	☐	☐	☐	☐	☐
Self-confidence (intellectual)	☐	☐	☐	☐	☐
Self-confidence (social)	☐	☐	☐	☐	☐

2. Overall high school GPA (best estimate)? _____

3. Did you take the following courses in high school, and if "yes," what grade did you receive?

	Yes	No	Grade
Biology	☐	☐	|A| |B| |C| |D| |F|
Chemistry	☐	☐	|A| |B| |C| |D| |F|
Physics	☐	☐	|A| |B| |C| |D| |F|
Physical Science	☐	☐	|A| |B| |C| |D| |F|
Algebra I	☐	☐	|A| |B| |C| |D| |F|
Algebra II	☐	☐	|A| |B| |C| |D| |F|
Geometry	☐	☐	|A| |B| |C| |D| |F|
Tigonometry	☐	☐	|A| |B| |C| |D| |F|
Calculus	☐	☐	|A| |B| |C| |D| |F|

4. Did you take the following courses at RCC (or equivalent at another college), and if "yes," what grade did you receive?

	Yes	No	Grade
Chemistry 2A	☐	☐	|A| |B| |C| |D| |F|
Chemistry 2B	☐	☐	|A| |B| |C| |D| |F|
Chemistry 3	☐	☐	|A| |B| |C| |D| |F|

5. How many college courses in the following areas have you completed here at RCC or at any other college?

	AT RCC					OTHER COLLEGES				
	0	1	2	3	4 or more	0	1	2	3	4 or more
Biological Sciences	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physics	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Science/Geology	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Algebra	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Geometry and/or Trigonometry	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Calculus	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

6. How many hours per week (outside of class time) do you expect to spend studying for this course? _____.

7 What letter grade do you expect or believe you will receive in this course? .....|A|..|B|..|C|..|D|..|F|

APPENDIX II

VARIABLES

The table in this appendix provides a comprehensive listing of the variables which were considered as possible candidates for submission to final regression analyses. The table provides the variable name used in the computerized analyses (first column); a description of the variable (second column); the type of variable (second-to-last column), that is, whether the variable was classified as continuous (C), dichotomous (D), or nominal (N); and finally (last column) either the range of values (for the C and D variables) or the number of categories (No. of Cats.) for the nominal variables. The column labeled "Source" (the third) specifies the origin of the variable: S1 (1996 Student Information Form), S2 (RCCD student records), T (California Chemistry Diagnostic Test), R1 (researcher-fashioned variable based on the source referenced above the "R"), and R2 (researcher-fashioned variable based on factor analysis and the source referenced above the "R"). All dichotomous variables were coded such that the higher of the two numbered codes represented the condition which the variable name implies; for instance, for the variable SEXFEM, a value of "2" signifies the condition "being female," while a "1" is the "otherwise" or "not" condition, in this case "being male."

To further assist the reader through the tables, the following brief discussion based on the variable SYNGPAB is presented. This dichotomous variable was constructed by the researcher using

information available on student's college GPA from RCCD's computer banks (S2). Construction of a dichotomous variable such as SYNGPAB is advantageous in that students without previous college experience, i.e. no cumulative GPA, need not be ignored or discounted during regression (treated as "missing values"), since the constructed variable now connotes two distinct classes of subjects: those who have a recorded cumulative college GPA equal to or greater than a 3.0 (the "2"s) and those who do not have a recorded cumulative college GPA of 3.0 or better (the "1"s). The variable effectively separates subjects into two levels of college experience: students with a history of college performance at or above the B-grade level and those without this history.

Table AII-1

Listing of Variables

Variable Name	Variable Description	Source	Type	Range or No. of Cats.
ACAB	Student self-rated academic ability	S1	C	1-5
AGE	Student age	S2	C	16-50
ALG#	Number of college courses previously taken in algebra	S1	C	0-4
ASIAN	Ethnicity Asian/Pacific Islander	S2	D	1-2

BALGTPS#	The sum of college courses previously taken in biology, geometry/trigonometry, algebra, and physical science (BS# + GT# + ALG# + PS#)	S1 R2	C	0-11
BLACK	Ethnicity Black/African American	S2	D	1-2
BS#	Number of college courses previously taken in biological sciences	S1	C	0-4
CALC#	Number of college courses previously taken in calculus	S1	C	0-4
CALPHYS#	The sum of college courses previously taken in calculus and physics (CALC# + PHYS#)	S1 R2	C	0-4
CHMGRD	Course grade for the semester in General Chemistry I	S2	C	0-4
DTEST	Student's score on the CCDT	T	C	2-37
ETHNIC	Student ethnic grouping	S2	N	5
EXPGRD	Grade student expects to receive in General Chemistry I	S1	C	0-4
EXSTUDY	Hours per week student expects to spend studying General Chemistry I	S1	C	0-40
FALLUN	Number of units student enrolled in for fall term '96 (units attempted for the fall)	S2	D	3-18
GENKNO	Student self-rated general knowledge	S1	C	1-5
GETBACH	Student's educational goal is the attainment of a baccalaureate degree	S2 R1	D	1-2
GOAL	Student's educational goal	S2	N	13
GT#	Number of college courses previously taken in geometry and trigonometry	S1	C	0-4
HASRCCGP	Student had a cumulative GPA on RCCD records	S2	D	1-2
HISPANIC	Ethnicity Hispanic	S2	D	1-2
HRSATT	Total units attempted as District student	S2	C	0-160
HRSCMP	Total units completed as District student	S2	C	0-134
HRSPAS	Total units passed as District student	S2	C	0-131

HSA1	Took Algebra I in high school	S1	D	0-1
HSA1GD	High Algebra I grade	S1	C	0-4
HSA1GDA	Student received an A in high school Algebra I	S1 R1	D	0-1
HSA1GDB	Student received a B or better in high school Algebra I	S1 R1	D	0-1
HSA1GDC	Student received a C or better in high school Algebra I	S1 R1	D	0-1
HSA2	Took Algebra II in high school	S1	D	0-1
HSA2GD	High Algebra II grade	S1	C	0-4
HSA2GDA	Student received an A in high school Algebra II	S1 R1	D	0-1
HSA2GDB	Student received a B or better in high school Algebra II	S1 R1	D	0-1
HSA2GDC	Student received a C or better in high school Algebra II	S1 R1	D	0-1
HSBFACT1	Total number of advanced high school science and mathematics classes where student achieved a B or better (HSCHMGDB + HSCLGDB + HSPHGDB + HSTRGDB)	S1 R2	C	0-4
HSBFACT2	Total number of fundamental high school science and mathematics classes where student achieved a B or better (HSA1GDB + HSA2GDB + HSBIOGDB + HSGMGDB)	S1 R2	C	0-4
HSBIO	Took biology in high school	S1	D	0-1
HSBIOGD	High school biology grade	S1	C	0-4
HSBIOGDA	Student received an A in high school biology	S1 R1	D	0-1
HSBIOGDB	Student received a B or better in high school biology	S1 R1	D	0-1
HSBIOGDC	Student received a C or better in high school biology	S1 R1	D	0-1
HSCHM	Took chemistry in high school	S1	D	0-1
HSCHMGD	High school chemistry grade	S1	C	0-4

HSCHMGDA	Student received an A in high school chemistry	S1 R1	D	0-1
HSCHMGDB	Student received a B or better in high school chemistry	S1 R1	D	0-1
HSCHMGDC	Student received a C or better in high school chemistry	S1 R1	D	0-1
HSCL	Took calculus in high school	S1	D	0-1
HSCLGD	High school calculus grade	S1	C	0-4
HSCLGDA	Student received an A in high school calculus	S1 R1	D	0-1
HSCLGDB	Student received a B or better in high school calculus	S1 R1	D	0-1
HSCLGDC	Student received a C or better in high school calculus	S1 R1	D	0-1
HSGGM	Took geometry in high school	S1	D	0-1
HSGMGD	High school geometry grade	S1	C	0-4
HSGMGDA	Student received an A in high school geometry	S1 R1	D	0-1
HSGMGDB	Student received a B or better in high school geometry	S1 R1	D	0-1
HSGMGDC	Student received a C or better in high school geomety	S1 R1	D	0-1
HSPH	Took physics in high school	S1	D	0-1
HSPHGD	High school physics grade	S1	C	0-4
HSPHGDA	Student received an A in high school physics	S1 R1	D	0-1
HSPHGDB	Student received a B or better in high school physics	S1 R1	D	0-1
HSPHGDC	Student received a C or better in high school physics	S1 R1	D	0-1
HSPS	Took physical science in high school	S1	D	0-1
HSPSGD	High school physical science grade	S1	C	0-4

HSPSGDA	Student received an A in high school physical science	S1 R1	D	0-1
HSPSGDB	Student received a B or better in high school physical science	S1 R1	D	0-1
HSPSGDC	Student received a C or better in high school physical science	S1 R1	D	0-1
HSTR	Took trigonometry in high school	S1	D	0-1
HSTRGD	High school trigonometry grade	S1	C	0-4
HSTRGDA	Student received an A in high school trigonometry	S1 R1	D	0-1
HSTRGDB	Student received a B or better in high school trigonometry	S1 R1	D	0-1
HSTRGDC	Student received a C or better in high school trigonometry	S1 R1	D	0-1
LNAB	Student self-rated lecture notetaking ability	S1	C	1-5
MJASCMTH	Student's declared major was in the natural sciences or mathematics	S2 R1	D	1-2
NIGHT	Student took General Chemistry I at night	S2 R1	D	1-2
PC2A	Took Chemistry 2A in college	S1	D	0-1
PC2AGD	Chemistry 2A grade	S1	C	0-4
PC2AGDA	Student received an A in Chemistry 2A	S1 R1	D	0-1
PC2AGDB	Student received a B or better in Chemistry 2A	S1 R1	D	0-1
PC2AGDC	Student received a C or better in Chemistry 2A	S1 R1	D	0-1
PC2BGD	Chemistry 2B grade	S1	C	0-4
PC2BGDA	Student received an A in Chemistry 2B	S1 R1	D	0-1
PC2BGDB	Student received a B or better in Chemistry 2B	S1 R1	D	0-1

PC2BGDC	Student received a C or better in Chemistry 2B	S1 R1	D	0-1
PC3	Took Chemistry 3 in college	S1	D	0-1
PC3GD	Chemistry 3 grade	S1	C	0-4
PC3GDA	Student received an A in Chemistry 3	S1 R1	D	0-1
PC3GDB	Student received a B or better in Chemistry 3	S1 R1	D	0-1
PC3GDC	Student received a C or better in Chemistry 3	S1 R1	D	0-1
PHYS#	Number of college courses previously taken in physics	S1	C	0-4
PS#	Number of college courses previously taken in physical science	S1	C	0-4
PSAB	Student self-rated problem solving ability	S1	C	1-5
QUANTAB	Student's overall score on self-rated quantitative abilities (ACAB + GENKNO + PSAB + SRAB + SLFCNFI)	S1 R2	C	12-25
RCCGPA	Student's cumulative GPA at RCCD	S2	C	0.00-4.00
RCCMAJ	Student's declared college major	S2	N	43
RDAB	Student self-rated reading ability	S1	C	1-5
SEXFEM	Student's sex is female	S1	D	1-2
SLFCNFI	Student self-rated self-confidence (intellectual)	S1	C	1-5
SLFCNFS	Student self-rated self-confidence (social)	S1	C	1-5
SRAB	Student self-rated science reasoning ability	S1	C	1-5
SUCC	Student successfully completed General Chemistry I	S2 R1	D	0-1
SYNGPAA	Has a cumulative college GPA >= 3.5	S2 R1	D	1-2

SYNGPAB	Has a cumulative college GPA >= 3.0	S2 R1	D	1-2
SYNGPAC	Has a cumulative college GPA >= 2.0	S2 R1	D	1-2
VERBAB	Student's overall score on self-rated verbal abilities (LNAB + RDAB + WRAB)	S1 R2	C	3-15
WHITE	Ethnicity White/Caucasian	S2 R1	D	1-2
WRAB	Student self-rated writing ability	S1	C	1-5
WRKHRS	Number of hours/week student works on a job	S2	C	0-60

It should be noted most of the variables described in the tables above were never considered in final computer runs—for reasons discussed in the analysis of the data—yet, they are still listed to assist the reader in following various arguments related to the variables. Additional variable details may be found in the main text and from inspection of the survey instrument (Appendix I).

APPENDIX III

MISSING VALUES, FREQUENCIES, and OTHER DATA HOUSEKEEPING

This appendix focuses a number of data housekeeping details that needed to be carried out before data could be submitted for final review.

In SPSS any case which has a "missing value" for a variable (denoted by a "." in the computer output) is not included in the analysis. In the present situation—with many variables involved—this would not only reduce statistical power but could also distort the data such that erroneous conclusions relative to the originally tested group could easily be reached rather than those more appropriate to the group. Hence, in the interest of being able to present a substantially accurate description of the relationship among the variables, the following procedures were developed and employed for the handling of missing values.

Preliminary frequency distributions, histograms, and descriptive statistics were run on the variables listed in Appendix II. Among statistics obtained were: mean, median, standard deviation, variance, skewness, minimum, maximum, and range. Space here does not permit a comprehensive listing of the frequency distributions and statistics for all the variables involved in the study[3] ; however, some numbers for a

---

[3] A preponderance of descriptive and correlational statistical information on variables was compiled during this project, not all of which has been placed in this document; these are available, however, from the author at the address listed among the references, upon request.

representative variable ACAB (students' self-rated academic ability) are displayed in Figure AIII-1 alongside a histogram of the ACAB data (contrasted against a normal curve). The histogram reveals the distribution of ACAB data to be somewhat normal.

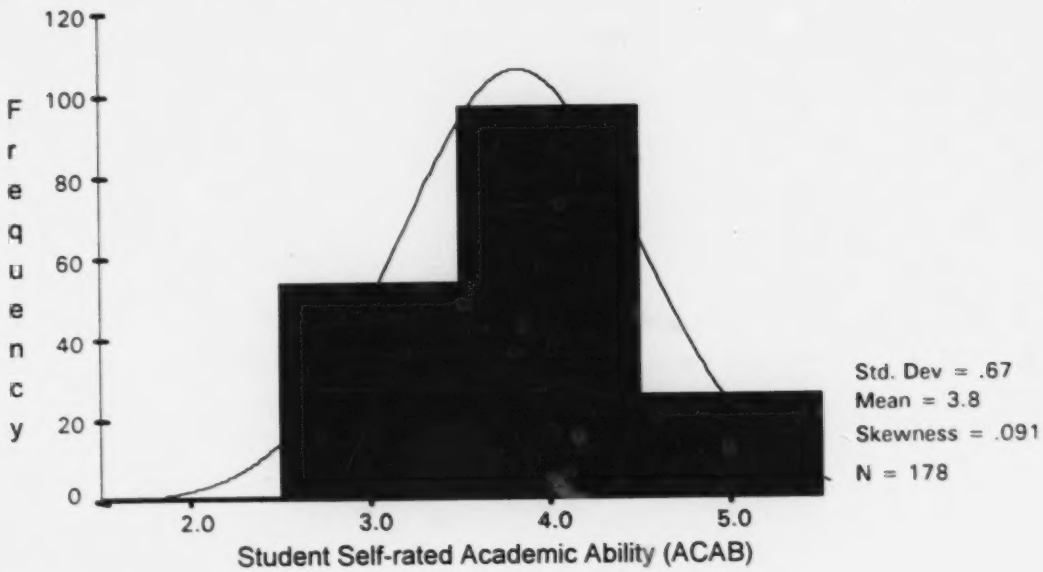

Figure AIII-1 Histogram of Academic Ability

Other tests of normality can be performed via examination of box plots, normal plots, detrended normal plots, skewness, and range/standard deviation (SPSS, 1992a; Glass & Hopkins, 1984; Hills, 1981). A few examples are in order: Skewness values equal to or greater than 2, or equal to or less than -2 should alert one to marked deviations from normal distributions which ideally have a skewness of zero (.091 in the figure). Another crude test of normality (see Hills, 1981) is to divide the range of scores by the standard deviation; for

BEST COPY AVAILABLE

normally distributed numbers and for samples ranging in size from 25 to 1000 subjects, these ratios should be between 4 and 6 (for ACAB the ratio is 4.49). Thus, the normality of the ACAB data is satisfied from a multiple perspectives criterion. Granted, regression procedures do not require that independent variables be normally distributed (Glass & Hopkins, 1984, p. 141); nevertheless, it is still useful to know from preliminary analyses which variables are normal and which are not. For the most part, variables in this study (except of course the dichotomies) were approximately normal. Notable exceptions were some of the variables which designated number of college courses previously taken—CALC#, GT#, PHYS#, and PS#. These had a definite positive skewness to their stature, simply because most of the 178 Chemistry 1A Ss had not taken college calculus (70%), or physics (82%), and so on. The distribution for CCDT scores (Figure AIII-2) was found to be quite normal; that for chemistry grades (Figure AIII-3) was skewed slightly to the left.

Returning once again to the discussion on missing values, correction procedures were applied as in Georgakakos (1995): Where it made sense, for certain continuous variables (e.g., high school GPA, expected hours of study), means were substituted; for others, such as self-rated abilities and dichotomous variables, the whole number closest to the mean was employed. For categorical variables, the mode was used (this only occurred for one case whose ethnicity was not known). Overall, corrections were small to nonexistent. For the entire general chemistry sample, the most missing values occurred for expected study

hours (2.8% of the cases) and high school GPA (2.2%); the replacement means were 9.0 weekly hours and 3.2. No correction to self-rated ability variables ever exceeded 1.1%; no correction to self-reported high school courses and grades variables ever exceeded 1.8%. All the above percentages should be viewed as upper limits to the missing value corrections, since if the truth be known about the missing values, they in all likelihood would not be too far away from the values supplied. Finally, no missing values occurred for any of the college course experience or grade variables (ALG#, GT#, CALC#, PC2A, PC2B, PC3, etc.). In short, corrections for missing values are nominal and non-perturbative.

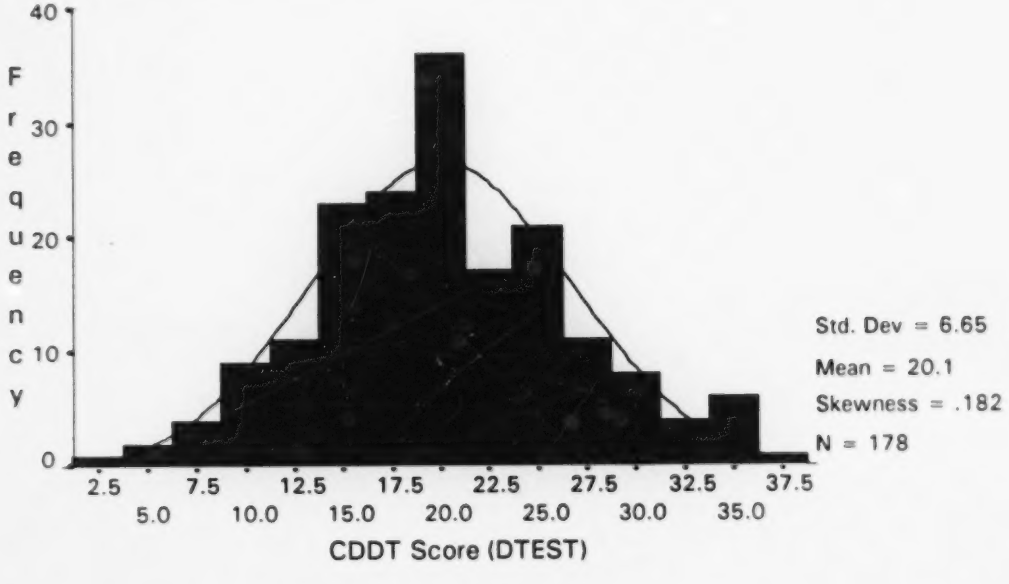

Figure AIII-2 Histogram of Student Scores on the California Chemistry Diagnostic Test

BEST COPY AVAILABLE

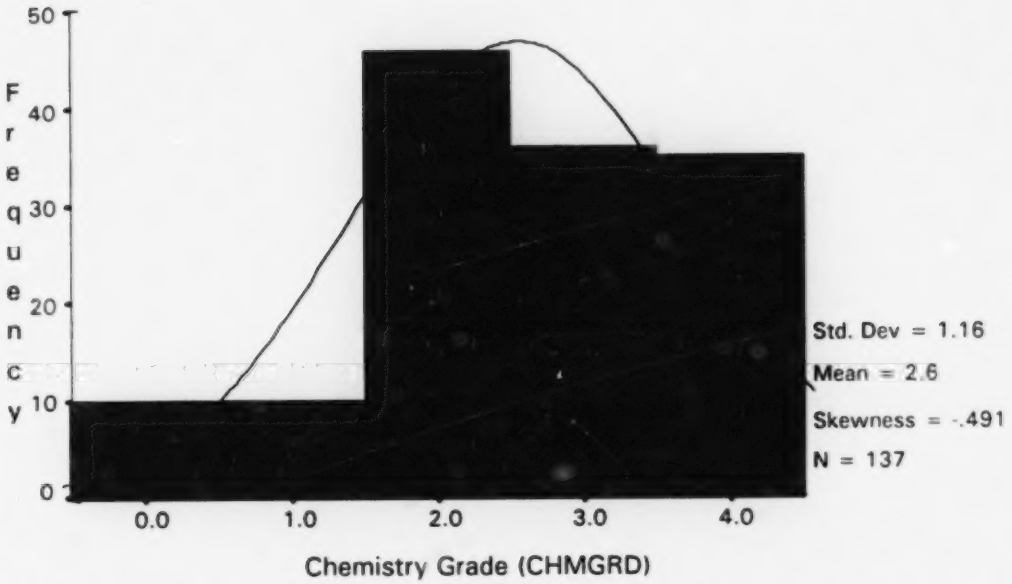

Figure AIII-3 Histogram of Grades in General Chemistry

Majors

When filling out the Riverside Community College Admission Application (RCCD, 1996), students select a major from a coded list of 112 possible majors. District majors data on Ss ultimately became distilled to the dichotomous variable (MJASCMTH), with one group consisting of science and mathematics majors and the other of students who chose from codes other than science or mathematics. Of 178 Ss, 48 were science or math majors (27 biology, 9 chemistry, 2 botany, 5 geology, 2 mathematics, 1 microbiology, 1 physics, 1 zoology).

This section has dealt with some of the details involved in getting the data ready for final processing, with focus primarily on how missing values were treated. It has been shown that corrections made to the data were not large, in general of the order of a few percent. Finally, it

BEST COPY AVAILABLE

should be noted that total corrections are not additive across variables and in fact are probably much less than the various mean percentages reported, owing both to the correction methodology itself and to the fact that different students tended to be involved in these corrections from one variable to the next.

APPENDIX IV

VARIABLE REDUCTION and SCREENING REGRESSORS

This appendix synopsizes the process of variable reduction and screening that produced the final slate of 19 possible predictor variables ultimately submitted to the SPSS stepwise regression procedures. The main regression programs employed are also displayed.

Data Reduction Using Principal Component Analysis

One technique that may be employed to consolidate the effects carried by a the number of variables which have moderately high correlations with each other is principal components (PC) analysis (SPSS, 1992c). According to Afifi and Azen (1979), the PC technique is useful for describing the dependence structure of a set of p response variables $X_1 \ldots X_p$ having a multivariate distribution and a mean vector $\mu^{p\times 1} = (\mu_1 \ldots \mu_p)'$ and a variance/covariance matrix $\Sigma^{p\times p}$. As is the case with regression, the X's themselves need not be normal; however, $\Sigma^{p\times p}$ must be nonsingular, that is, its determinant must not be equal to zero (Afifi and Azen, 1979, p. 318; Tatsuoka, 1988, p. 146). Assuming the above conditions are met, one can then derive a set of transformed, uncorrelated (orthogonal) Y_p principal components which are unique linear combinations of X_1 to X_p (i.e., $Y_1 = \sum_{j=1}^{p} a_{1j} X_j$, ..., $Y_p = \sum_{j=1}^{p} a_{pj} X_j$) and whose variances capture all the variance represented by $\Sigma^{p\times p}$, but in their own transformed diagonalized "lambda" (Λ) matrix such that the diagonal elements of Λ (called eigenvalues λ_i) are the variances of Y_1 to

Y_p in descending order of magnitude, i.e., $\lambda_1 > \lambda_2 > \cdots \lambda_p$.[4] A commonly excepted practice in reducing the dimensionality of representation is to focus attention only on those Y components or factors whose λ values are equal to or greater than 1. Additional insights into, and hence ability to interpret, the nature of the relationship of the response variables to selected factors can be achieved by the structure-simplifying follow-up procedure known as factor rotation. A number of factor rotation algorithms are available (see Nunnally, 1978; Afifi & Clark, 1984); the current data reductions employed two (SPSS, 1992c): 1) orthogonal rotation (varimax procedure) which preserves the orthogonality of factors, and 2) oblique rotation (oblimin procedure) which allows for examination of rotated factors which might not be independent from each other.

Self-rated Ability Scales

Table AIV-1 presents the correlations among the 9 self-rated ability variables—ACAB, GENKNO, LNAB, PSAB, RDAB, SLFCNFI, SLFCNFS, SRAB, and WRAB—for each general chemistry sample. The

---

[4] If $\Sigma^{p \times p}$ is singular (determinant = 0), that signals a condition of multicollinearity within the set of response variables, a condition where such high correlations exist between certain of the X's that an independent set of Y's can not be established, since one or more of the response variables is simply functioning as a carbon copy of some linear combination of the remaining X's. The number of collinearities (or near collinearities) among predictors equals the number of λ's with values of 0 (or near 0). The solution to the dilemma is to remove the echo, as it were, by reducing the rank of $\Sigma^{p \times p}$.

moderately high and significant r values make this group of variables an excellent candidate for the data reduction procedure just described.

Table AIV-1

Correlations: Self-rated Abilities for N = 178 Sample (below diagonal) and N = 137 Sample (above diagonal)

		Variable (Represented as a Variable Number—V#)								
V#	Variable	1	2	3	4	5	6	7	8	9
1	ACAB	1.00	.55	.45	.59	.46	.56	.28	.45	.44
2	GENKNO	.53	1.00	.45	.60	.42	.53	.43	.50	.47
3	LNAB	.46	.46	1.00	.40	.57	.47	.32	.30	.58
4	PSAB	.60	.60	.34	1.00	.43	.58	.43	.64	.36
5	RDAB	.47	.43	.54	.46	1.00	.47	.28	.38	.76
6	SLFCNFI	.52	.53	.45	.54	.45	1.00	.49	.42	.39
7	SLFCNFS	.28	.44	.32	.38	.26	.51	1.00	.30	.32
8	SRAB	.44	.56	.29	.66	.40	.46	.28	1.00	.36
9	WRAB	.49	.48	.58	.39	.75	.38	.30	.38	1.00

Note. All correlations significant with p (2-tailed) ≤ .001.

PC analysis followed by varimax rotation produced a two-factor solution with variables ACAB, GENKNO, PSAB, SRAB, SLFCNFI, and SLFCNFS loading on Factor 1 and LNAB, RDAB, and WRAB loading on Factor 2 (regardless of which chemistry sample was used). Now, even though the social and intellectual self-confidence variables correlated .5 with each other (V#'s 6 and 7 in the table), earlier research on these variables exposed profound differences in the way they related to academic achievement and student performance on a standardized test of science reasoning ability (Georgakakos, 1995). Thus, despite their moderate correlation with each other, SLFCNFI and SLFCNFS are not of

a kind and should not be included in the same factor. This, coupled with the fact that the present r between social self-confidence ratings and chemistry grades is negligibly small (.018), prompted the decision not to include SLFCNFS in any further analyses. A second PC analysis was then executed on the remaining 8 self-rated ability variables; these results are captured in Figure AIV-1 for the N = 178 sample (a similar result pertained for the N = 137 group). Still, a two-factor model is prevalent. The pioneer PC analysis conducted in 1933 by Hotelling in educational testing (see Afifi & Clark, 1984, p. 310) uncovered two major components to student responses on examinations which tested their abilities: one was quantitative, the other verbal. Such an interpretation of factors appears warranted by the results in Figure AIV-1, where skills requiring quantitative competence, such as problem solving ability and science reasoning ability, are seen to load on Factor 1, while skills requiring verbal competency—reading, writing, and lecture notetaking ability—load primarily on Factor 2. And while intellectual self-confidence, general knowledge, and academic ability load slightly less on the quantitative aspect and a slightly more on the verbal aspect than do problem solving and science reasoning (which is not unsurprising), results still dictate placing the former variables in the quantitative camp. Armed with the above detail, this researcher then returned to his 1991 data collected on a college-wide sample of 843 Ss who had rated themselves on the same abilities (except for lecture notetaking ability) using the same rating scale (Georgakakos, 1995) and conducted a PC analysis on the ability variables used in that study to see how those

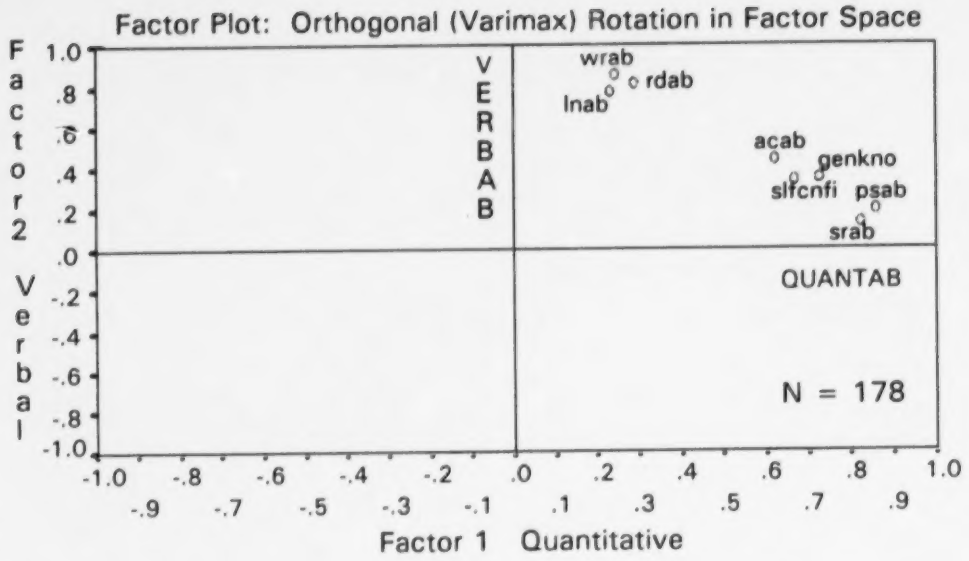

Figure AIV-1 Principal Components for Self-rated Abilities

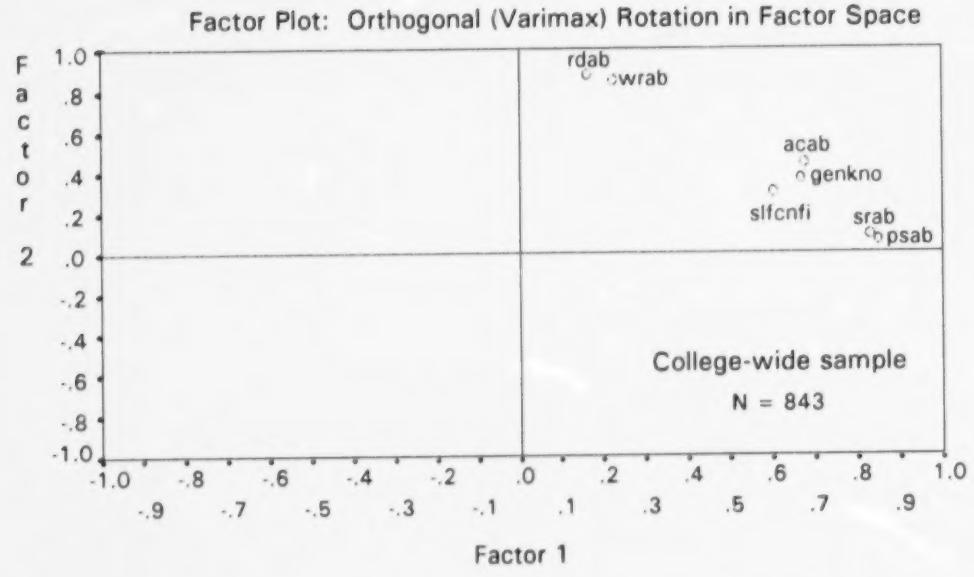

Figure AIV-2 Principal Components for Self-rated Abilities
Using 1991 College-wide Student Data

rotated factor solutions would compare with the ones in Figure AIV-1. The result is summarized by the two-factor model displayed in Figure AIV-2, very similar to the one above it and supporting the reduction of the eight present ability variables to two factor score variables. As stated in Afifi and Clark (1984, p. 350), when variables load highly on a factor, they may be given equal weighting. This makes the factor score simply the sum of the individual variable scores.[5] Thus, two new variables QUANTAB (student's overall score on self-rated quantitative abilities) and VERBAB (student's overall score on self-rated verbal abilities) were created—viz., QUANTAB = ACAB + GENKNO + PSAB + SRAB + SLFCNFI, and VERBAB = LNAB + RDAB + WRAB.

Background in College Science and Mathematics

Submitted to the same PC procedures as above were the 6 variables which spoke to number of college courses previously taken in science or mathematics: BS# (biological sciences), PHYS# (physics), PS# (physical science), ALG# (algebra), GT# (geometry and trigonometry), and CALC# (calculus). Another two-factor solution resulted (see Figure AIV-3). Clearly, experience in the more advanced courses—physics and calculus—loads almost exclusively on one factor,

---

[5] While other approaches to the construction of factor score equations are also justifiable on theoretical grounds (see Gorsuch, 1974), the tack taken here makes sense from a practical point of view as well, since it affords future researchers the ability of being able to convert, "on the spot," ability ratings such as above into overall verbal or quantitative scores

while experience in more basic science and math courses—biology, physical science, algebra, geometry and trigonometry—loads on the other. This is not surprising when one examines the various simple correlations involved. For example for the N = 178 sample, CALC# had a correlation of .43 with PHYS#, but had negative correlations of -.16 with ALG#, -.03 with BS#, -.02 with GT#, and -.18 with PS#. On the other hand PS# had positive correlations of .27 with ALG#, .18 with BS#, and .13 with GT#. (Virtually identical relationships exist for the N = 137 sample.)

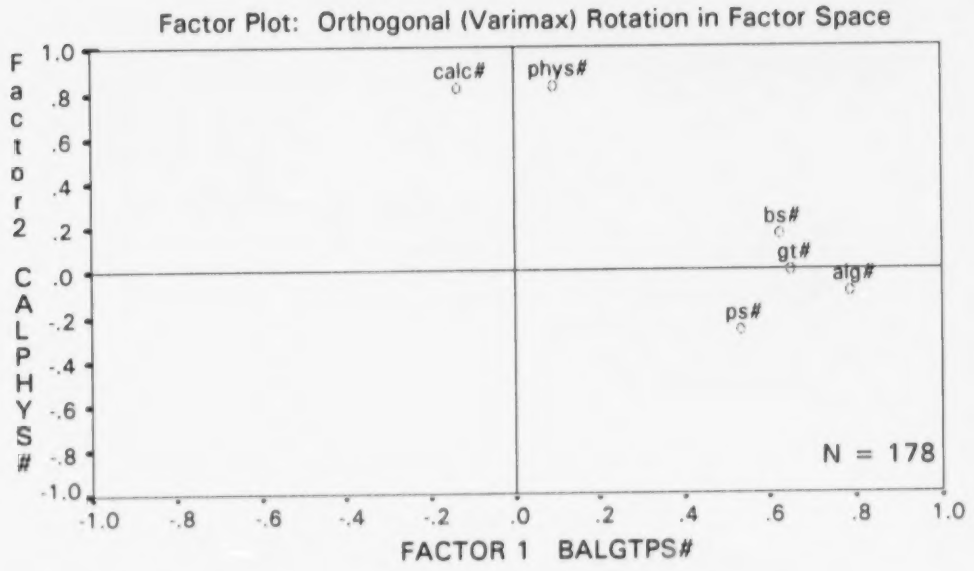

Figure AIV-3 Principal Components for College Courses Taken

It appears students' college science/math experience came from the more basic or from the more advanced courses but not from both.

Based on these findings two new variables were constructed: BALGTPS# = BS# + GT# + ALG# + PS# (i.e., total number of college courses previously taken in biology, geometry/trigonometry, algebra, and physical science) and CALPHYS# = CALC# + PHYS# (sum of college courses previously taken in calculus and physics). Six variables became reduced to two.

High School Course Grades

A preliminary PC run was carried out on the 27 dichotomous high school course grade variables (namely, grade = A, grade ≥ B, grade ≥ C for Algebra I, Algebra II, biology, chemistry, calculus, geometry, physics, physical science, and trigonometry—see Appendix II). Results demonstrated that these dichotomous variables tend to load on factors coursewise, not gradewise. That is, HSCHMGDA, HSCHMGDB, HSCHMGDC on one factor; HSA1GDA, HSA1GDB, HSA1GDC on another, and so on. A retrospective look at the zero-order correlation matrix for these variables (not given here) adds understanding, since the A, B, C variables for a given course correlate on average between .3 to .7 with each other but much less than this when compared to the grade variables from other courses. To illustrate for high school biology (N = 178 sample), receiving A correlated .46 with receiving a B or better, and .32 with receiving C or better; at the same time, B or better correlated .68 with C or better. In contrast, the highest correlation receiving C or better in biology had with any of the other high school course dichotomous grade variables was .21 with receiving C or better in

chemistry. Thus, much of the same information relative to a course is contained in the grade variables for a given course. In a comprehensive earlier study which examined the ability of dichotomous grade variables to act as predictors of performance for a variety of different samples (Georgakakos, 1995), it was found that grade-A or grade-B-or-better variables would sometimes be significant predictors, but never the grade-C-or-better variables. In the present exploration, across the spectrum of high school courses, the B-or-better variables had higher correlations with students' general chemistry course grades than did either of the other two forms of dichotomous grade variables. All of the above plus the additional fact that the B grade variables had inherently less skewness—each category more evenly represented—than the A dichotomies vouched for allowing the B-or-better variables to pass on to the next stage of analysis while the other 18 grade variables were summarily retired from further service.

This next stage included a factor analysis of the nine B-or-better high school grade variables. A three-factor solution resulted with physics, chemistry, calculus, and trigonometry loading on Factor 1; Algebra 1, Algebra 2, biology, and geometry on Factor 2; and physical science on Factor 3. Because the correlation between B or better in high school physical science and Chemistry 1A grades was only .03, and because the physical science variable was associated with the weakest of the three factors, the variable HSPSGDB was forced to the sidelines and a two-factor model was pursued using the 8 remaining B-or-better variables. Consequently, a new PC analysis followed by an oblique

rotation led to the factor structure shown in Figure AIV-4. The result is for the N = 178 group; that obtained for the N = 137 group was however essentially the same.

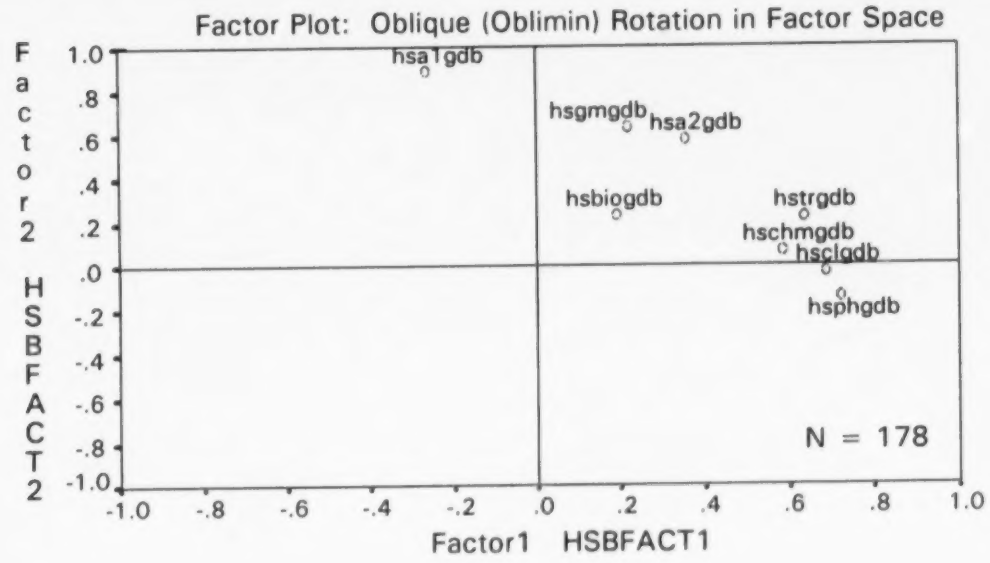

Figure AIV-4 High School Course Background Factors 1 & 2

That a more interpretable structure was achieved via an oblique rotation speaks to the correlation between the factors—good students tend to be good students. But the fact that courses project as they do, reminiscent of the college course results, speaks to the two tracks of achievement in science and math—the one advanced, the other more fundamental. In short, two new high school background variables emerged: **High school background factor 1** ... HSBFACT1 = HSCHMGDB + HSCLGDB + HSPHGDB + HSTRGDB ... the number of advanced high school science and mathematics classes where the

student achieved a B or better and **high school background factor 2** ... HSBFACT2 = HSA1GDB + HSA2GDB + HSBIOGDB + HSGMGDB ... the number of fundamental high school science and mathematics classes where the student achieved a B or better. Thus, eight variables were reduced to two.

Eliminations of Other Variables

Students' District experience in terms of units attempted (HRSATT), completed (HRSCMP), and passed (HRSPAS) had such high correlations with each other ($\geq .97$) that they in essence could serve as proxies for each other. The variable chosen to represent this aspect of a student's background was HRSPAS.

All Chemistry 2B background variables were scratched since only 4% of the Chemistry 1A students had taken this course, leaving the Chemistry 2B variables all abnormally skewed.

For the entire general chemistry sample the following correlations were obtained between college course grade A and B-or-better variables: $r_{SYNGPAA,SYNGPAB} = .55$, $r_{PC2AGDA,PC2AGDB} = .55$, and $r_{PC3GDA,PC3GDB} = .65$ (all with two-tailed $p < .001$). Since the previously discussed high school course grades were represented best by B grade variables, since the B-or-better variables are less skewed than the A cluster, and because the above correlations are relatively high—bringing to mind concerns about multicollinearity in regression—the variables SYNGPAA, PC2AGDA, and PC3GDA were relegated to the sidelines while the B variables were allowed to pass muster.

Other variables which were eliminated based on low correlations with prospective dependents (generally in the -.05 to .05 range) were EXPSTUDY, FALLUN, GETBACH, HRSPAS, MJASCMTH, and WRKHRS. For instance, the variable GETBACH (goal of attaining a bachelor's degree) had no significant correlation with any of the host of variables examined for the entire general chemistry sample. Now, although certain other variables such as SEXFEM, HISPANIC, BALGTPS#, and QUANTAB could have been excluded using this same criterion, they were not. Instead they were included so that one might get a general feeling for the types of interrelationships involved from an examination of the final correlation matrix presented in Appendix V.

Final List of Regression Variables

The following list of 19 independent variables was submitted to the regression analyses: AGE, ASIAN, BALGTPS#, BLACK, CALPHYS#, DTEST, EXPGRD, HISPANIC, HSBFACT1, HSBFACT2, HSGPA, NIGHT, PC2AGDB, PC3GDB, QUANTAB, SEXFEM, SYNGPAB, VERBAB, WHITE. The ratio of cases to variables for the entire chemistry sample is 178/19 = 9.4; for the sample used to predict grades this ratio is reduced to 137/19 = 7.2.

SPSS Regression Programs

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL selection
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT chmgrd
 /METHOD=FORWARD age asian balgtps# black calphys# dtest expgrd hispanic
 hsbfact1 hsbfact2 hsgpa night pc2agdb pc3gdb quantab sexfem syngpab verbab
 white .


LOGISTIC REGRESSION succ
 /METHOD=FSTEP(COND) age asian balgtps# black calphys# dtest expgrd
 hispanic hsbfact1 hsbfact2 hsgpa night pc2agdb pc3gdb quantab sexfem
 syngpab verbab white
 /SAVE PRED PGROUP COOK LEVER DFBETA RESID LRESID SRESID ZRESID
DEV
 /CLASSPLOT /CASEWISE OUTLIER(2)
 /PRINT=CORR ITER(1)
 /CRITERIA PIN(.05) POUT(.10) ITERATE(20) .
```

APPENDIX V

CORRELATIONS AND BASIC STATISTICS

Table AV-1 presents means and standard deviations for final regression variables for each chemistry sample. Table AV-2 (spread over two pages) gives correlations for these same 21 variables.

Table AV-1

Basic Statistics: Chemistry Grades, Success and 19 Potential Predictors

Variable	Mean (N=137)	Std Dev (N=137)	Mean (N=178)	Std Dev (N=178)
CHMGRD	2.56	1.16	N/A	N/A
SUCC	N/A	N/A	.66	.48
AGE	23.88	6.36	23.99	6.35
ASIAN	1.20	.40	1.19	.39
BALGTPS#	3.14	2.50	3.12	2.49
BLACK	1.03	.17	1.05	.21
CALPHYS#	.91	1.52	.85	1.50
DTEST	20.76	6.48	20.08	6.65
EXPGRD	3.54	.52	3.53	.54
HISPANIC	1.12	.33	1.14	.34
HSBFACT1	1.42	1.22	1.38	1.24
HSBFACT2	2.86	1.23	2.80	1.22
HSGPA	3.21	.50	3.20	.47
NIGHT	1.16	.37	1.17	.38
PC2AGDB	.45	.50	.46	.50
PC3GDB	.12	.32	.10	.30
QUANTAB	18.69	2.77	18.61	2.78
SEXFEM	1.47	.50	1.48	.50
SYNGPAB	1.37	.49	1.34	.48
VERBAB	10.80	2.30	10.83	2.25
WHITE	1.55	.50	1.54	.50

Table AV-2

Correlations: Chemistry Grades, Success, and Independent Variables for N = 178 Sample (**below diagonal**) and N = 137 Sample (**above diagonal**)

		Variable (Represented as a Variable Number—V#)									
V#	Variable	0	1	2	3	4	5	6	7	8	9
0	CHMGRD→ SUCC ↘	**1.00**	.18*	-.21*	-.03	-.20*	.12	.27**	.18*	-.07	.21*
1	AGE	.08	**1.00**	-.16	.34**	-.09	.10	-.03	.13	-.04	-.21*
2	ASIAN	-.08	-.14	**1.00**	-.26**	-.09	.03	-.10	.05	-.19*	.16
3	BALGTPS#	.04	.33**	-.28**	**1.00**	.13	-.11	-.23**	-.04	-.03	-.32**
4	BLACK	-.19*	-.03	-.10	.13	**1.00**	-.02	-.08	-.10	-.07.	-.06
5	CALPHYS#	.11	.08	.01	-.07	-.05	**1.00**	.20*	.02	.26**	.23**
6	DTEST	.23**	-.09	-.07	-.21**	-.08	.20**	**1.00**	.18*	-.10	.19*
7	EXPGRD	.06	.04	-.04	-.02	-.11	.04	.21**	**1.00**	-.18*	.13
8	HISPANIC	-.03	-.09	-.19*	-.06	-.09	.21**	-.08	-.15	**1.00**	.05
9	HSBFACT1	.15*	-.23**	.17*	-.31**	-.09	.28**	.26**	.14	.08	**1.00**
10	HSBFACT2	.18*	-.14	.15*	-.35**	-.14	.21**	.19*	-.01	.09	.50**
11	HSGPA	.12	-.07	-.06	-.30**	-.05	.07	.20**	.09	.09	.39**
12	NIGHT	.02	.28**	.01	-.01	-.03	.07	.05	.07	.08	.08
13	PC2AGDB	.04	.22**	-.18*	.25**	.02	.00	.16*	.16*	-.13	-.22**
14	PC3GDB	.12	-.05	-.02	.17*	.02	-.04	-.13	-.12	.09	-.09
15	QUANTAB	.06	.11	-.16*	.07	-.01	.09	.21**	.37**	-.06	.21**
16	SEXFEM	-.01	.01	-.06	.14	.01	-.20**	-.13	-.04	.05	-.11
17	SYNGPAB	.25**	.31**	-.10	.28**	-.16*	.04	.01	.21**	-.01	-.11
18	VERBAB	.05	.19*	-.31**	.25**	.05	-.06	.06	.12	-.02	.03
19	WHITE	.09	.22**	-.52**	.15	-.24**	-.10	.21**	.10	-.43**	-.18*

(Continued on following page)

BEST COPY AVAILABLE

Table AV-2 (continued)

Correlations: Chemistry Grades, Success, and Independent Variables for N = 178 Sample (**below diagonal**) and N = 137 Sample (**above diagonal**)

V#	Variable	Variable (Represented as a Variable Number—V#)									
		10	11	12	13	14	15	16	17	18	19
0	CHMGRD↘ SUCC ↓	.20*	.22**	.15	.17	.00	.09	-.07	.30**	.14	.21*
1	AGE	-.12	-.06	.23**	.15	-.06	.19*	.05	.30**	.24**	.23**
2	ASIAN	.12	-.08	-.07	-.19*	-.01	-.23**	-.10	-.08	-.39**	-.55**
3	BALGTPS#	-.35**	-.34**	-.01	.23**	.18*	.09	.18*	.29**	.29**	.13
4	BLACK	-.02	-.01	-.08	.02	.07	-.06	.10	-.13	.05	-.19*
5	CALPHYS#	.20*	.07	.12	-.02	-.04	.06	-.14	.09	-.06	-.13
6	DTEST	.15	.18*	.02	.24**	-.15	.12	-.17*	.00	.02	.26**
7	EXPGRD	-.03	.09	.04	.17*	-.07	.33**	-.05	.25**	.08	.13
8	HISPANIC	.06	.02	.20*	-.11	.07	-.04	.00	.03	-.07	-.41**
9	HSBFACT1	.47**	.36**	.11	-.19*	-.09	.11	-.11	-.07	-.05	-.15
10	HSBFACT2	**1.00**	.50**	.11	-.15	-.09	.07	-.01	.05	-.02	-.10
11	HSGPA	.50**	**1.00**	.06	-.10	.01	.14	.15	.11	.18*	.05
12	NIGHT	.10	.08	**1.00**	-.11	-.04	.16	-.05	-.01	.10	.04
13	PC2AGDB	-.18*	-.11	-.06	**1.00**	-.28**	-.00	.13	.34**	.08	.17
14	PC3GDB	-.08	.00	-.06	-.27**	**1.00**	-.02	.07	.00	.03	.01
15	QUANTAB	.12	.18*	.15*	-.03	-.06	**1.00**	-.20*	.11	.60**	.27**
16	SEXFEM	-.05	.17*	.00	.09	.09	-.18*	**1.00**	.16	.15	.03
17	SYNGPAB	.00	.11	.01	.39**	-.01	.04	.13	**1.00**	.17*	.03
18	VERBAB	.01	.24**	.15*	.02	.03	.61**	.18*	.12	**1.00**	.28**
19	WHITE	-.09	-.00	.04	.17*	.01	.18*	.01	.10	.19*	**1.00**

Note. A -.00 value means a near zero yet slightly negative correlation,

* signifies $p \leq .05$; ** signifies $p \leq .01$ (2-tailed).

BEST COPY AVAILABLE

APPENDIX VI

ADJUSTED R^2 AND MALLOW'S C_p

For a regression equation based on a sample of size N, involving p added independent variables which describe or predict some criterion in a larger population, the adjusted multiple correlation coefficient, $\overline{R}$, may be computed from the multiple correlation coefficient, R, with use of an equation given in Afifi and Clark (1984):

$$\overline{R}^2 = R^2 - \frac{p(1-R^2)}{N-p-1} . \qquad \text{Eq (A6-1)}$$

The nature of Equation (A6-1) is to furnish $\overline{R}^2$ which increase to some maximum value then decrease as more independent variables are added to the regression. R^2, on the other hand, always increases as p increases and also tends to overestimate population values. The equation, then, acts to compensate for a sampling bias that might ascribe too much importance to minor predictors. Some researchers suggest variable selection be halted when $\overline{R}^2$ reaches a maximum.

Another criterion that may be used is Mallow's C_p statistic (Afifi & Clark, 1984):

$$C_p = (N - p - 1)\left[\frac{MSE(p)}{MSE(k)} - 1\right] + p + 1, \qquad \text{Eq (A6-2)}$$

where MSE(p) and MSE(k) are the residual mean squares based on entry of p and k (total) variables, respectively. As more variables are selected, the residual sum of squares SSE(p) decreases an eventually becomes equal to SSE(k), making the square-bracketed term tend to

zero; however, the p + 1 term increases. A selection strategy based on C_p, then, will stop adding variables at the p which minimizes C_p.

It will be recalled that the partial *F* test was the selection/stopping rule of choice for all linear regressions. Once variables could no longer satisfy the *F* requirement at the $\alpha = .05$ level, then regression was halted. Study of all regressions performed on all samples revealed that $\overline{R}^2$ steadily increased and C_p steadily decreased right on through addition of the last variable to pass the *F* test. This finding is encouraging and helps establish confidence in the models: Absence of reversals in either $\overline{R}^2$ or C_p may be taken as an indication that no extraneous or "unneeded" predictor variables were entered into the regression equations.

For that matter, raising the alpha level of the *F* test to, say, .1 would increase statistical power and also allow for insertion of more predictors. However, in the interest of parsimony, a desire to avoid Type I errors, a wish to focus on only the most important predictors—all arguments for smaller models noted at the beginning of the Analysis of Data and Results section—as well as in keeping with the more-traditionally employed variable-selection procedures (Astin, 1991), the partial *F* test was given the last word. Although the final models selected in the study—using the procedures and strategies described here and elsewhere—may not be the "best" of all possible models, the diagnostic statistical computations performed on them point to the fact that the models that were ultimately arrived at are reasonably <u>good</u> ones (see also, Kleinbaum, Kupper, & Muller, 1988, p. 318).

APPENDIX VII

ANALYSIS OF MODELS: SPLIT SAMPLES AND CROSS-VALIDATION

Introduction

One method of assessing the reliability of a model is to conduct a split-sample analysis (Kleinbaum, Kupper, & Muller, 1988), whereby each member of a given sample is randomly assigned to one of two groups, and separate regression analyses are performed on each subgroup. A qualitative check of reliability can be made from a comparison of the predictors obtained in the various regressions—total sample with subgroup 1, with subgroup 2, etc.. Equivalent predictors across all groups would be indicative of high reliability.

A more quantitative check can be made by looking at the "shrinkage on cross-validation," (Shr). Generally speaking, let $R^2(1)$ be the square of the correlation between subgroup 1's observed and predicted dependent variable scores, that is, the squared multiple correlation coefficient, $r^2(Y_1, \hat{Y}_1)$. If the regression equation obtained for subgroup 1 is now used to predict outcome values for subgroup 2 (denoted as $\hat{Y}_2^*$), and these predicted scores are, in turn, correlated with the observed scores for subgroup 2, the square of the correlation which results, $r^2(Y_2, \hat{Y}_2^*)$—called the "cross-validation correlation" and symbolized as $R_*^2(2)$—can be subtracted from $R^2(1)$ to obtain the shrinkage:

$$Shr = R^2(1) - R_*^2(2). \qquad \text{Eq (A7-1)}$$

Kleinbaum, Kupper, and Muller (1988, p. 330) note that shrinkage values less than .10 are indicative of a reliable model, while those greater than .90 clarion an unreliable one.

Split-sample Analyses on N = 137 Group: Chemistry Grades

The chemistry sample was first stratified on the basis of sex, ethnicity, and CCDT scores to help insure a greater conformability between selected subsamples. The computer was then programmed to randomly select two distinct, nonoverlapping subsamples of approximately equal size—subgroup 1 (N = 68) and subgroup 2 (N = 69). A total of eight subsamples—four pairings—were selected. The linear regression program in Appendix IV was then applied, yielding regression equations for each subgroup. The most commonly entered variables were those that were entered for the undivided sample—i.e., SYNGPAB, DTEST, HSBFACT1, and ASIAN (or WHITE). Other predictors, such as, AGE, NIGHT, PC2AGDB, and HSBFACT2 (each once) and BLACK (twice), however, showed up as significant predictors in certain subsample equations, exposing an inconstancy among predictors. Equations and shrinkage calculations for one of the pairings are given here: Subgroup 1's equation (CHMGRD = .744 + .624 × WHITE + .618 × SYNGPAB) led to $R^2(1) - R_*^2(2) = (.3849)^2 - (.3213)^2 = .045$ ... meaning, subgroup 1's equation predicts almost as well for subgroup 2 as it does for its own group. Subgroup 2's equation (CHMGRD = -.685 + .0485 × DTEST + 1.107 × SYNGPAB + .454 × HSBFACT1) led to $R^2(2) - R_*^2(1) = (.664)^2 - (.171)^2 = .41$ ... a much less favorable result, i.e., the equation predicts quite well for its own

subgroup but has problems forecasting for subgroup 1. The eight shrinkage values computed over all the subsamples studied ranged from .035 (considered good reliability) to .41 (considered dubious), with a mean Shr = .19. It should be noted for all split-sample regressions conducted that $\overline{R}^2$ and Mallow's C_p behaved as with the full sample—the former increasing steadily, the latter decreasing always, until regression was halted on the basis of the partial *F* tests. True, $\overline{R}^2$, C_p, collinearity diagnostics, and residual analyses all uphold the reliability of the full-sample model, yet the split-samples analysis tends to abrade this support with mixed results. Much of this skittishness is because halving the full sample also halves the ratio of cases to variables, from 7.2 to 3.6. This dip below the suggested safety limit of 5 puts a strain on the smaller samples' (not necessarily the full sample's) capacity to yield consistent equations. Further testimony to this claim is found by comparing standard errors for regression coefficients of variables that predict in both full- and half-sample equations: b for SYNGPAB in the final full-sample equation was .71, compared with .68 for this variable in a final split-sample equation. While these b's are similar, their standard errors, on the other hand, are not: .18 in the N = 137 sample versus .27 in the N = 68 sample—a 50% increase in error.

In sum, the present split-sample results can only give lukewarm support to the full-sample regression equation. However, in light of the reasons discussed above, one should be more inclined to place less emphasis on this reliability criterion and more on others.

Split-sample Analyses on N = 178 Group: Success

Split-sample/cross-validation data obtained relative to the full sample (N = 178) linear regression where SUCC was the criterion were less vagarious and more respectable than those discussed above, primarily because half samples here were 30% larger (89 vs. 68). Ten shrinkage values from five different subgroup pairs ranged from .04 to .24 and had a mean of .11. Variables selected to subsample final equations were mostly those chosen for the full sample—SYNGPAB, DTEST, PC3GDB, and HSBFACT2. Others selected were HSBFACT1 (once), BLACK (twice), and QUANTAB (once).

Another difference between the analysis here and the one above is that for like variables, the standard errors for regression coefficients in relation to the sizes of the b's themselves are not too different in the split-half sample than in the full sample (see Table AVII-1 below).

Table AVII-1

Comparison of Final Step Regression Constants and Standard Errors

	Full Sample (N = 178)		Half Sample #1 (N = 89)		Half Sample #2 (N = 89)	
Variable	b	Std Err b	b	Std Err b	b	Std Err b
SYNGPAB	.246	.070	.237	.093	.269	.102
DTEST	.016	.005	.023	.007	.023	.007
PC3GDB	.262	.111	.237	.093	.329	.144
HSBFACT2	.060	.028	did not enter		did not enter	
constant	-.185	.155	-.135	.201	-.195	.197

Briefly then, support for the reliability of the full-sample dichotomous linear regression equation from the present split-sample runs is more forward, less lackadaisical than that established for the prediction of grades.

APPENDIX VIII

ANALYSIS OF MODELS: RESIDUALS

Background on Residuals

For a linear regression based on a sample of N cases and k predictors—i.e., $N-k-1$ degrees of freedom (df), the standardized and studentized residuals for the *i* th case may be written, respectively, as: $z_i = e_i/S$ and $r_i = z_i/\sqrt{1-h_i}$, where e_i, the unstandardized residual, is the difference between the observed and predicted dependent variable and expressed as $e_i = Y_i - \hat{Y}_i$; S is the estimator of the population standard deviation expressed as $S = \sqrt{\sum_{i=1}^{N} e_i^2 \Big/ (N-k-1)}$; and h_i is the leverage, a measure of the importance of the *i* th observation in determining the model fit, with values such that $0 \leq h \leq 1$ (Kleinbaum, Kupper, & Muller, 1988). Since h_i helps take into account the tendency of the *i* th observation to "mask" its own variance, absolute values for r_i will be equal to or greater than the corresponding z_i values. Furthermore, while the studentized residuals approximately follow a Student's *t* distribution, their behavior can be expected to closely coincide with that of the standardized residuals—e.g., standard normal (mean 0, variance 1) distribution—as degrees of freedom increase above 30. In the present case—chemistry sample for grade prediction with N = 137, k = 4, and hence df = 132—one should expect both z_i and r_i to follow the normal distribution, each having only about 5% of the examinees with residual absolute values in excess of 1.96. Moreover, as the size of the sample

increases, the probability of a single case being able to influence or "distort" the regression model is commensurately reduced. This claim is examined briefly below.

Testing Residuals for Outliers

One approach suggested by Kleinbaum, Kupper, and Muller (1988, p. 199) for testing whether a particular residual is significantly different from zero (has an extreme value not due to chance alone) is to use a corrected or family type-I error rate. For example, using the chemistry sample and a two-tailed test with an original significance level $\alpha = .05$, one may correct for chance rejections by dividing .025 by 132 which gives the appropriate p-rejection value (family α) of 1.9×10^{-4}, which in turn requires $|r_i|$ to be ≥ 3.55. No residual in the chemistry-grades sample came close to this value; the studentized residual of the most distant outlier was -2.47 (standardized -2.39).

Another measure of the influence of the i^{th} observation is Cook's distance (d_i), a computation of how much the regression coefficients are changed by deletion of the particular observation in question. For a special case where all dependent variables (X_j 's) have equal variances, a mean of 0, and are uncorrelated, d_i is proportional to $\sum_{j=1}^{k}[\hat{\beta}_j - \hat{\beta}_{j(-i)}]^2$, where $\hat{\beta}_j$ is the estimated regression coefficient of X_j using all the data, and $\hat{\beta}_{j(-i)}$ is this estimate when case i is withheld. Kleinbaum, Kupper, and Muller (1988) also give a more general expression for d_i, in terms of leverages and studentized residuals, applicable to any set of data:

$$d_i = \left(\frac{1}{k+1}\right) r_i^2 \left(\frac{h_i}{1-h_i}\right). \qquad \text{Eq (A8-1)}$$

From the definition of r_i given earlier in the section, one may obtain an expression for the leverage in terms of z_i and r_i ... that is:

$$h_i = 1 - \left(\frac{z_i}{r_i}\right)^2. \qquad \text{Eq (A8-2)}$$

Inserting the extreme chemistry sample residuals given above into Equation (A8-2) gives $h_i = .064$ (a relatively small number); d_i may then be computed as:

$$d_i = \left(\frac{1}{4+1}\right)(-2.47)^2\left(\frac{.064}{.936}\right) = .083.$$

Cook and Weisberg (1982) suggest $d_i > 1$ deserve close scrutiny. No individual observation in the N = 137 sample had a value for d_i greater than .083. In fact, this was the largest Cook's distance for any observation in either of the two linear regressions carried out. Thus, all observations were retained in the analyses of each group.

Because the degrees of freedom involved for each sample studied was large, residual analyses based on either z_i or r_i yielded virtually identical results. Henceforth—at least until the last section of this appendix when logistic regression diagnostics is taken up—all residual reportings refer only to standardized or unstandardized residuals.

Linear Regressions:

Most of the following pages summarize the residual analyses related to the linear regression equations established to predict grades

BEST COPY AVAILABLE

and chance for success in general chemistry. The analyses attempt to speak directly to the assumptions of linear regression stipulated earlier: 1) normal distribution of residuals (normality of Y's), 2) linearity between Y's and $\hat{Y}$'s, and 3) homogeneity of residual variances. Discussion begins with the N =137 sample and continues with the N = 178 sample. Residuals obtained for the loglinear regression are dealt with in the final pages, and a comparison between the linear and logit approaches is made.

Grades as Dependent Variable (N = 137)

A distribution of the residuals associated with the regression equation used for the prediction of chemistry grades is presented in Figure AVIII-1. There is accord between the distribution and the superimposed normal curve. A standard normal distribution should

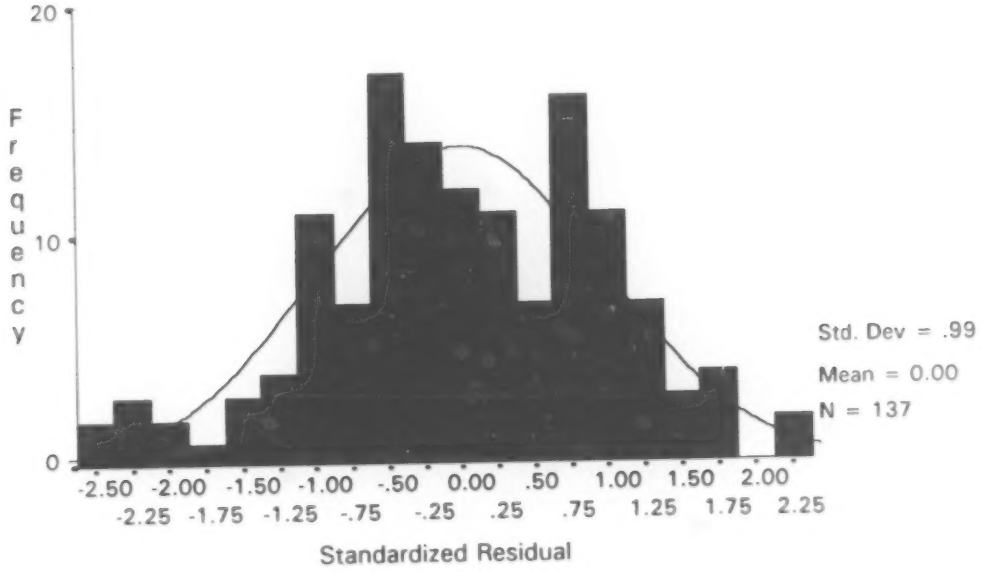

Figure AVIII-1 Histogram of Standardized Residuals

BEST COPY AVAILABLE

have 4.56% and .27% of the residuals with absolute z scores in excess of 2 and 3, respectively. The actual percentages of cases fitting into these two categories for this sample are 6.6% (9 cases) and 0%, quite close to normal probability expectations. An excessive number of cases with $|z_i| \geq 2$ would be a flag for potential problems. However, here no outliers needed to be discarded. Further verification of the normality assumption may be found in a normal probability plot (Figure AVIII-2), where the relationship between the cumulative percentage of standardized residuals and that expected for a normal distribution of residuals is most linear (hovering around the reference line).

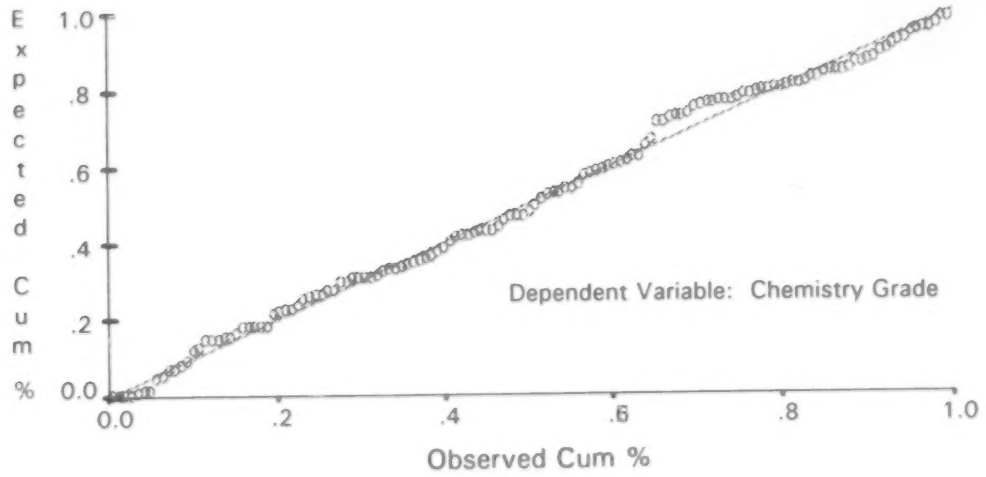

Figure AVIII-2 Normal Plot of Regression Standardized Residuals

Furthermore, as should be, the correlation between the residuals and predicted grades was zero, and the residuals had no correlation with any of the predictor variables.

Plots of standardized residuals versus SYNGPAB and CCDT scores (Figures AVIII-3 & 4) bear witness to the homoscedasticity of these predictors. Similar plots were found for ASIAN and HSBFACT1. (Figure AVIII-3 can also be compared with the partial residual plot[6] of grades vs. SYNGPAB in Figure AVIII-5, where a positive slope reflective of the positive regression coefficient for this variable is easily observable.)

Linearity was also observed in a plot (not shown here) of chemistry grades versus predicted chemistry grades. Hence, the present residual analysis results support the conclusion that none of the regression assumptions—normality, linearity, homoscedasticity—was violated.

Linear Regression: SUCC as Dependent Variable (N = 178)

Residual analyses in linear regressions which employ dichotomous dependent variables are not as "clean" as when the

---

[6] The partial residual plot, also called a partial regression plot, presents an interesting alternative perspective on the regression results (SPSS, 1992a). Here residuals computed for the dependent variable when it is predicted from all independent variables excluding the independent variable of interest are plotted against residuals for the given independent when it is predicted from all the other independent variables. The net effect is to remove the linear effects of the other independent variables from the criterion and the independent variable of concern.

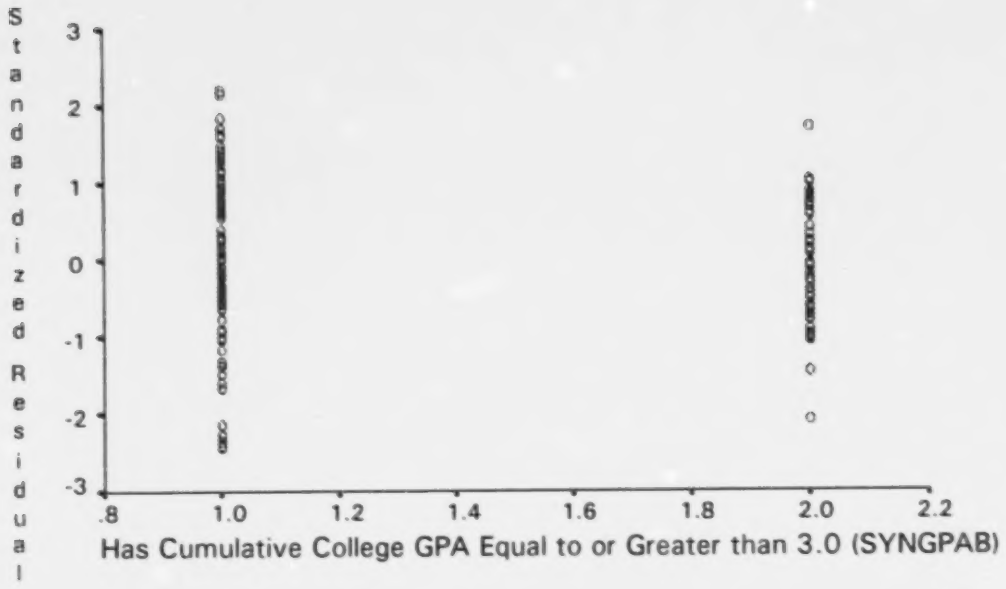

Figure AVIII-3 Z_i vs. SYNGPAB for Chemistry-grades Sample

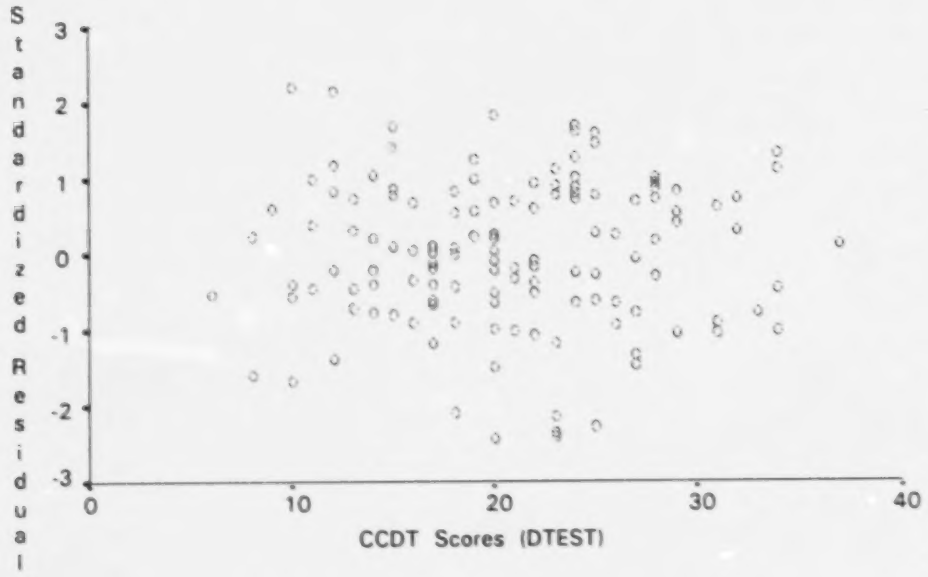

Figure AVIII-4 Z_i vs. CCDT Scores for Chemistry-grades Sample

BEST COPY AVAILABLE

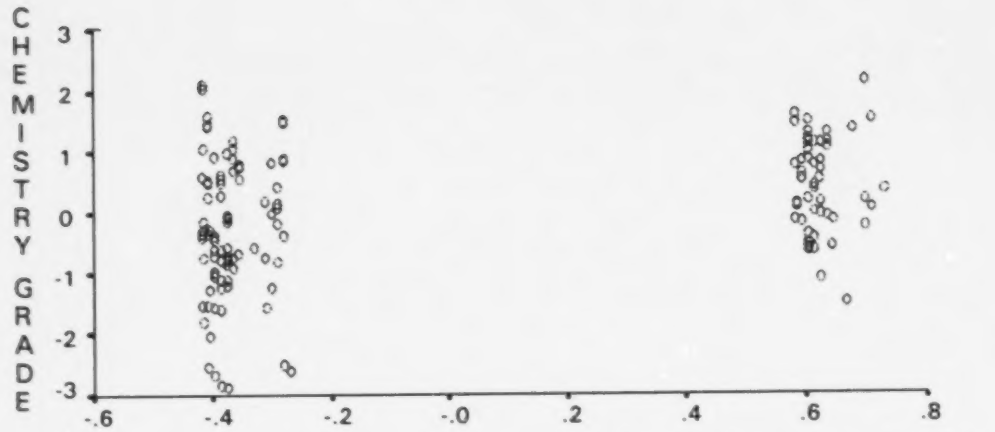

Figure AVIII-5 Partial Regression Plot: Partial Residuals of CHMGRD vs. Partial Residuals of SYNGPAB

dependent variables are continuous. Part of the problem is that predicted values for success are stored in the computer as continuous numbers; these ranged from of .16 to 1.25 (incidently, the range of logit probabilities, .14 to .98, is attributable to the same two Ss). Residuals are then computed by subtracting these continuous prediction scores from the actual success values (the 0's and 1's of the variable SUCC), leaving behind a bimodal distribution of residuals as shown in Figure AVIII-6. Misclassifications occur for the high and low values of z, but not the in-between values. To illustrate, consider a student who was unsuccessful in completing the course (SUCC = 0) but has a linear regression predicted value for success of .87; his residual will be -.87

BEST COPY AVAILABLE

(standardized residual = -1.98, for this data). This student will be classified a false positive. By the same token a successful student with a low predicted score, say, .25 will have a residual of .75 (z_i = 1.70) and thus become a false negative.

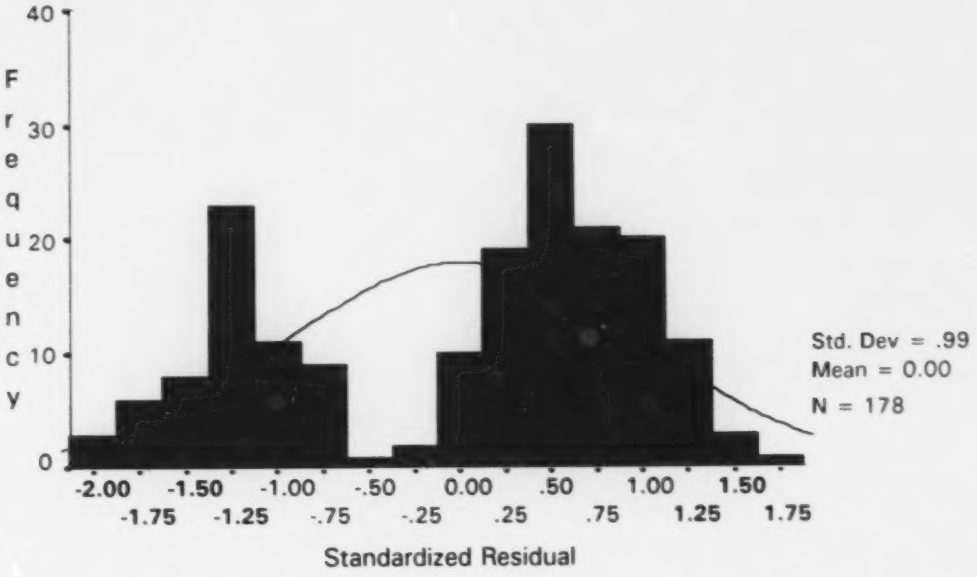

Figure AVIII-6 Distribution of Standardized Residuals

From the graph, the three bars to the extreme right contain the 13 false negatives (note: because of graphical scaling procedures the third band from the extreme right contains 11 cases—2 of whom are correctly placed, 9 of whom are not; the upper two bands contain the remaining 4 false negatives). The 39 false positives are contained in the four leftmost bars with 1 case correctly classified sharing the **" -1.25 bar"** with 22 false positive cases. Before leaving these residuals, it should be pointed out that they still have a mean of zero, and have a zero

BEST COPY AVAILABLE

correlation with the predicted success scores. None of their absolute values is aberrantly large. In fact no Cook's distance for any participant in this regression exceeded .04.

One way in which residual results may be "smoothed out" or simplified is to integerize them by rounding predicted success scores; thus, values < .5 become 0, while those ≥ .5 become 1. This removes the bimodal character of the residual frequency distribution, yielding the distribution shown in Figure AVIII-7. Notice how the misclassified Ss become instantly discernible on either side of the central band (correctly classified Ss); also notice that the histogram tends to look more "normal."

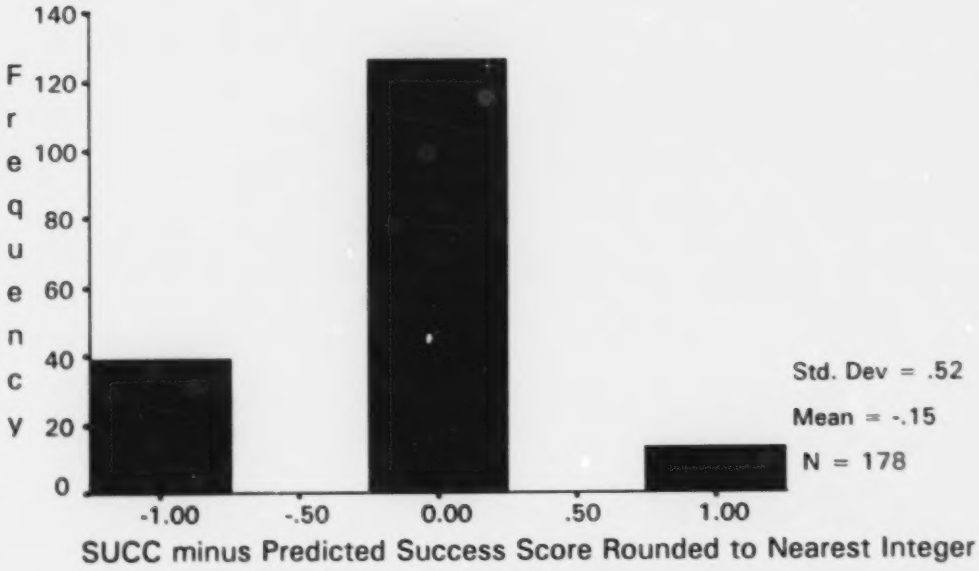

Figure AVIII-7 Histogram of Integerized Residuals

BEST COPY AVAILABLE

Remaining residual plots for this sample need not be displayed since their basic message has already been broadcast through the corresponding figures presented for the chemistry-grades sample. Basic conclusions, also, are as earlier: model reliability supported by the post-regression data.

Logistic Regression Diagnostics (N = 178)

While SPSS computes standardized residuals $z_i = e_i / \sqrt{P_{S_i}(1-P_{S_i})}$, where e_i is the difference between the observed and model-predicted probabilities for success for the i^{th} subject, a somewhat less cumbersome treatment of aberrant data and hence useful handle on model reliability can be found from a close scrutiny of the deviances and studentized residuals computed and saved during the logistic regression diagnostics routine performed by SPSS. Individual deviances DEV_i are calculated using the algorithm $DEV_i = \pm\sqrt{-2\ln(P_j)_i}$, where P_j is the calculated probability of a subject being in the group he actually belongs to (i.e., either P_{S_i} or $1-P_{S_i}$); the minus sign before the radical being involved if the predicted group is not the same as the actual group. Like the standardized residuals, the deviances should be approximately normally distributed with a mean of zero and a unit standard deviation. Such was the case in the present investigation, although some bimodality due inevitably to the dichotomous nature of the outcome variable was again in evidence. What is more important is that no case had a deviance whose absolute value equaled or exceeded 2.0, which again means no single case was an inordinate influence on the logistic

regression results. This is further attested to by the fact that the mean Cook's distance over all cases was only .03, with no single value greater than .27.

In the SPSS logit diagnostics, an individual's studentized residual, again called r_i, is simply that deviance which is computed for a case if that case is withheld from the regression analysis. Large divergences between DEV_i and r_i should be cause for concern. Figure AVIII-8 is a plot of $(DEV_i - r_i)$ versus Case#. Notice, most of the differences are minimal, and with the maximum difference being only about .09, no single case should be considered an outlier.

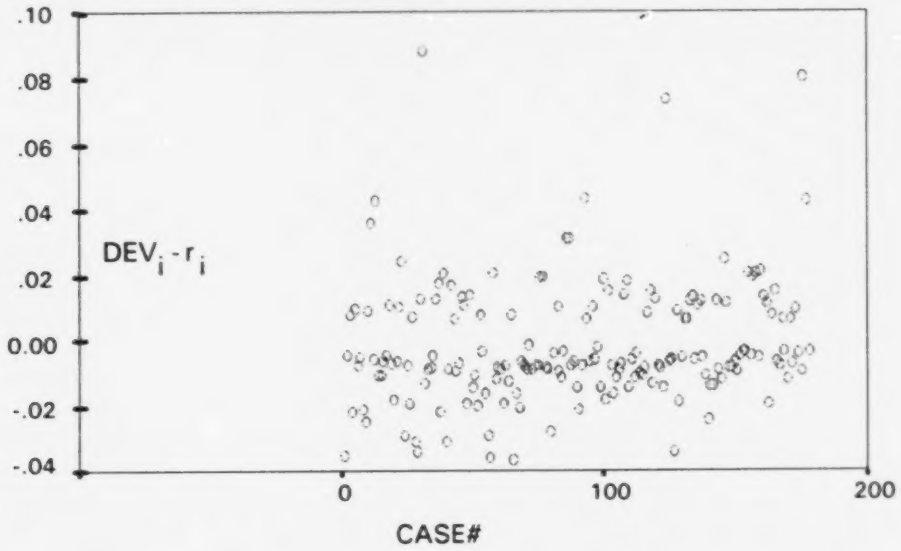

Figure AVIII-8 Residual Differences for Logistic Regression

Finally, the table below compares some basic statistics for $DEV_i - r_i$ with a comparable test statistic applicable to linear regression, $z_i - r_i$. Again, maximum differences between standardized and studentized residuals are small, regardless of regression method.

Table AVIII-1

Basic Statistics for Residual Differences from Different Regression Methods

Regression Method	N	Variable	Mean	Std Dev	Min	Max
Logistic	178	$DEV_i - r_i$	0	.02	-.036	.089
Linear Dichotomous	178	$z_i - r_i$	0	.01	-.038	.049
Linear Continuous	137	$z_i - r_i$	0	.02	-.052	.080

In sum, the residual analysis undertaken here may be used in support of the reliability of the loglinear model.

APPENDIX IX

SOME COMPARISONS OF GENERAL CHEMISTRY RESULTS

CCDT Score Means and General Chemistry Course Grades

Table AIX-1 displays the overall *F* results for an analysis of variance (ANOVA) where CCDT score is the dependent and chemistry grade is the grouping variable. The significant main-effect *F* of 3.122 (p = .017) indicates that significant differences should be found between one of more pair of test score means.

Table IX-1

ANOVA Table for Chemistry 1A CCDT Scores with Chemistry Grades as Factor

Source	SS	DF	MS	*F*	Sig of *F*
Main Effect	492.98	4	123.24	3.122	.017
CHMGRD					
Explained	492.98	4	123.24	3.122	.017
Residual	5210.07	132	39.47	1.94	
Total	5703.05	136	41.93		

For multiple *t* tests (c contrasts), each employing significance level α, the overall probability α_Σ (also called the family error rate or family alpha) of making a Type I error may be expressed as: $\alpha_\Sigma = 1 - (1 - \alpha)^c$. Hence, the chance of committing a Type I error increases dramatically as the number of comparisons made increases. To safeguard against reporting false positives one may lower the value of the per-contrast α

used to establish whether null hypothesis H_0 should be rejected or retained. For an "acceptable" α_Σ an applicable pairwise α may be obtained using the relationship: $\alpha = (\alpha_\Sigma/c)$. This approach—referred to as the least significant difference (LSD) or Bonferroni method (Kleinbaum & Kupper, 1978, p. 265; Kleinbaum, Kupper, & Muller, 1988)—is quite conservative and consequently helps insure against the occurrence of an inordinate number of false rejections of H_0.

Table IX-2

Pairwise Contrasts of CCDT Score Means by General Chemistry Course Grade

Chemistry Grade (N)	CCDT Mean	t-test p (1-tailed) for H_0: Mean (i) = Mean (j)				
		i \ j	A	B	C	D
A (35)	23.71	A	.			
B (36)	20.78	B	.030	.		
C (46)	19.56	C	.002*	.389	.	
D (10)	18.10	D	.010*	.127	.245	.
F (10)	18.50	F	.015	.162	.305	.445
Overall (N = 137)	20.76					

* Significant at the $\alpha_\Sigma/c = .01$ level.

Pairwise contrasts results are summarized in Table IX-2 for the different course-grade CCDT means. The maximum number of contrasts for five groups is $(5 \times 4 \div 2)$ or 10; α_Σ was set equal to .1, so only contrasts whose *t* tests were significant at an α level of approximately .01 became significant. The null hypothesis tested is: the mean for group i = the mean for group j, where i and j are dummy indices (here

represented by the letter grades themselves). The *t*-test probabilities are the off-diagonal elements of the i-j crossings. The most striking contrasts are between the CCDT "A" mean (23.71) and the "C" and "D" means (19.56 and 18.10). The positive relationship between the CCDT and course grades is evident.

10 Factor Comparison of Successful with Unsuccessful Ss

An appropriate statistical comparison of two groups on a number of variables first requires establishing a statistical difference between the groups on the basis of the variables overall before proceeding to examine the groups on a variable-by-variable basis. The omnibus procedure is called a multivariate analysis of variance (MANOVA). It may serve as preamble to an ANOVA, just as the ANOVA is precursor to individual *t* tests (Hummel & Sligo, 1971; Barker & Barker, 1984). To test the equality of several means with two independent samples one may invoke the Hotelling T^2 statistic:

$$T^2 = \frac{N_1 N_2}{N_1 + N_2}(\overline{X}_1 - \overline{X}_2)' S^{-1}(\overline{X}_1 - \overline{X}_2), \qquad \text{Eq (A9-1)}$$

where $\overline{X}_1$ and $\overline{X}_2$ are the vectors of the means for groups 1 and 2, and S^{-1} is the inverse of the pooled within-groups covariance matrix (Afifi & Azen, 1979). A significant Hotelling T^2 indicates the two groups are dissimilar with respect to the dependent variable vectors.

A MANOVA was performed, examining the mean vectors associated with the 10 input variables displayed in Table IX-3; the grouping variable was SUCC. The value computed for Hotelling's T^2 (using 10 hypothesis and 167 error degrees of freedom) was 4.23, very

significant ($p < .001$). The two groups could then be compared on each variable whose mean and standard deviation are given in the table. Dichotomous variable means reflect group percentages: for example, for Asians in the unsuccessful and successful categories, the values are 23.0% and 16.2%, respectively, with the difference in these percentages not statistically significant. Statistical significance is observed, however, for differences between the subgroup means for CCDT score, being Black, high school background factors 1 & 2, and the college GPA variable.

Table IX-3

Multiple Variable Contrasting: Unsuccessful with Successful Categories of General Chemistry using the Entire Sample for Reference

	SUCC Value and (N)						
	0 (N = 61)		1 (N = 117)			Entire Sample (N =178)	
Variable	Mean	Std Dev	Mean	Std Dev	2-tailed p	Mean	Std Dev
*DTEST	17.967	6.578	21.180	6.443	.002	20.079	6.649
ASIAN	1.230	.424	1.162	.370	.277	1.185	.390
*BLACK	1.098	.300	1.017	.130	.047	1.045	.208
HISPANIC	1.148	.358	1.128	.336	.722	1.135	.343
WHITE	1.475	.504	1.573	.497	.219	1.539	.500
SEXFEM	1.492	.504	1.479	.502	.868	1.483	.501
*HSBFACT1	1.115	1.279	1.513	1.208	.042	1.376	1.244
*HSBFACT2	2.492	1.273	2.957	1.163	.015	2.798	1.218
PC3GDB	.049	.218	.128	.336	.098	.101	.302
*SYNGPAB	1.180	.388	1.427	.497	.001	1.343	.476

* Precedes a variable whose difference between subgroup means is statistically significant ($p \leq .05$).

REFERENCES

Afifi, A. A., & Azen, S. P. (1979). Statistical analysis: A computer oriented approach (2nd ed.). New York: Academic Press.

Afifi, A. A., & Clark, V. (1984). Computer-aided multivariate analysis. New York: Van Nostrand Reinhold Co.

Aldrich, J. H., & Nelson, F. D. (1984). Linear probability, logit, and probit models. Beverly Hills, CA: Sage Publications.

Astin, A. W. (1987). Achieving educational excellence. San Francisco: Jossey-Bass.

Astin, A. W. (1991). Assessment for excellence. New York: Macmillan.

Astin, A. W., Christian, C. E., & Henson, J. W. (1975). The impact of student financial aid programs on student choice. Los Angeles: Higher Education Research Institute.

Astin, A. W., Green, K. C., Korn, W. S., Schalit, M., & Berz, E. (1988). The American freshman: National norms for fall 1988. Los Angeles: Higher Education Research Institute.

Barker, H. R., & Barker, B. M. (1984). Multivariate analysis of variance (MANOVA): A practical guide to its use in scientific decision making. Tuscaloosa, AL: University of Alabama Press.

Board of Governors California Community Colleges (1993). Title V matriculation regulations. Sacramento, CA: Board of Governors.

Boruch, R. F., & Creager, J. A. (1972). Measurement error in social and educational survey research (ACE Research Reports Volume 7, No. 2.). Washington, DC: American Council on Education.

BouJaoude, S. B., & Giuliano, F. J. (1991). The relationship between students' approaches to studying, formal reasoning ability, prior knowledge, and gender and their achievement in chemistry. Syracuse, NY: Paper presented at the Annual Meetings of the National Association for Research in Science Teaching, Lake Geneva, WI. (ERIC Document Reproduction Service No. ED 332 877)

Chandran, S., Treagust, D. F., & Tobin, K. (1987). The role of cognitive factors in chemistry achievement. Journal of Research in Science Teaching, 24 (2), 145-160.

Cook, R. D., & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart, & Winston.

Dickinson, D. J., & O'Connell, D. Q. (1990). Effect of quality and quantity of study on student grades. Journal of Educational Research, 83 (4), 227-231.

Garcia, T., Yu, S. L., & Coppola, B. P. (1993). Women and minorities in science: Motivational and cognitive correlates of achievement. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA, April. (ERIC Document Reproduction Service No. ED 359 235)

Georgakakos, J. H. (1990). Prediction of success and grades in political science and history: A matriculation study. Riverside, CA: Research study done at Riverside Community College. (ERIC Document Reproduction Service No. ED 317 228)

Georgakakos, J. H. (1995). Impact of differential college environments on the science reasoning ability of community college students: A matriculation study. Doctoral dissertation, School of Education, University of California, Riverside.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Gorsuch, R. L. (1974). Factor analysis. Philadelphia: Saunders.

Hills, J. R. (1981). Measurement and evaluation in the classroom (2nd ed.). Columbus, OH: Charles E. Merrill Publishing Co.

Hummel, T. J., & Sligo, R. J. (1971). Empirical comparisons of univariate and multivariate analysis of variance procedures. Psychological Bulletin, 76, 49-57.

Johnson, T. F., & Butts, D. R. (1983). The relationship among college science student achievement, engaged time, and personal characteristics. Journal of Research in Science Teaching, 20 (4), 357-366.

Karpp, E. R. (1995). Paths to success, volume III. Validating the California Chemistry Diagnostic Test for local use. Glendale, CA: Research study done at Glendale Community College. (ERIC Document Reproduction Service No. ED 384 402)

Kleinbaum, D. G., Kupper, L. L., & Muller, K. E. (1988). Applied regression analysis and other multivariate methods (2nd ed.). Boston: PWS-Kent Publishing Co.

McMillan, J. H., & Schumacher, S. (1984). Research in education: A conceptual introduction. Boston: Little Brown & Co.

Nordstrom, B. H. (1989). Predicting performance in freshman chemistry. Research report. (ERIC Document Reproduction Service No. ED 347 065)

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Pace, C. R. (1982). Achievement and the quality of student effort. Washington, DC: Department of Education. Paper presented at a meeting of the National Commission on Excellence in Education, May. (ERIC Document Reproduction Service No. ED 227 101)

Pascarella, E. T., & Terenzini, P. T. (1991). How college affects students. San Francisco: Jossey-Bass.

Razali, S. N. & Yager, R. E. (1994). What college chemistry instructors and high school chemistry teachers perceive as important for incoming college students. Journal of Research in Science Teaching, 31 (7), 735-747.

Riverside Community College District. (1994). Riverside Community College matriculation appeals petition. Riverside, CA: RCCD.

Riverside Community College District (1996). Riverside Community College admission application. Riverside, CA: RCCD.

Riverside Community College District. (1997). Riverside Community College schedule of classes: Spring 1997. Riverside, CA: RCCD.

RCCD Office of Planning and Development. (1996). RCCD factbook. Riverside, CA: RCCD.

Russell, A. A. (1994). A rationally designed general chemistry diagnostic test. Journal of Chemical Education, 71(4), 314-317.

Spencer, H. E. (1996). Mathematical SAT test scores and college chemistry grades. Journal of Chemical Education, 73(12), 1150-1153.

SPSS (1988). SPSS-X user's guide (3rd ed.). Chicago: SPSS Inc.

SPSS (1992a). SPSS for Windows: Base system user's guide, release 5.0. Chicago: SPSS Inc.

SPSS (1992b). SPSS for Windows: Advanced statistics, release 5.0. Chicago: SPSS Inc.

SPSS (1992c). SPSS for Windows: Professional statistics, release 5.0. Chicago: SPSS Inc.

Tatsuoka, M. M. (1988). Multivariate analysis (2nd ed.). New York: Macmillan.